IRE Transactions on ELECTRONIC COMPUTERS

Volume EC-8 **JUNE, 1959** **Number 2**

Published Quarterly

TABLE OF CONTENTS

7882
5I2

PUBLISHED BY THE
Professional Group on ELECTRONIC COMPUTERS

IRE PROFESSIONAL GROUP ON ELECTRONIC COMPUTERS

The Professional Group on Electronic Computers is an association of IRE members with professional interest in the field of Electronic Computers. All IRE members are eligible for membership, and will receive all Group publications upon payment of a fee of $2.00 per year, 1959.

PGEC OFFICERS

RICHARD O. ENDRES, *Chairman*

ARNOLD A. COHEN, *Vice-Chairman* WILLIAM J. BARTIK, *Secretary-Treasurer*

PGEC ADMINISTRATIVE COMMITTEE

Term Ending 1960	*Term Ending 1961*	*Term Ending 1962*
W. L. EVANS	W. L. ANDERSON	Y. M. HILL
R. F. JOHNSTON	D. HAAGENS	E. C. JOHNSON
S. R. OLSON	F. E. HEART	L. G. F. JONES
C. W. ROSENTHAL	K. W. UNCAPHER	R. W. MELVILLE
W. H. WARE		M. J. RELIS

COMMITTEES

Membership
R. T. BLAKELY, *Chairman*
T. O. ELLIS, *Vice-Chairman for Affiliates*

Awards
A. A. COHEN, *Chairman*

Student Activities
R. W. MELVILLE, *Chairman*

Sectional Activities
W. S. SPEER, *Chairman*

Publications
L. C. HOBBS, *Chairman*

Bibliography
L. G. F. JONES, *Chairman*

Foreign Relations
W. H. WARE, *Chairman*

Long-Range Planning
A. A. COHEN, *Chairman*

PGEC EDITORIAL BOARD

H. E. TOMPKINS, *Editor*

J. R. ANDERSON
E. J. MCCLUSKEY, JR.
J. E. ROBERTSON
M. RUBINOFF
J. R. WEINER

IRE Transactions® on Electronic Computers

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Electronic Computers at 1 East 79th Street, New York 21, N.Y. Responsibility for the contents rests upon the authors and not upon the IRE, the Group, or its members. Price per copy: IRE-PGEC members, $2.15; IRE members, $3.20, nonmembers, $6.45. Yearly subscriptions rate: nonmembers, $17.00; colleges and public libraries, $12.75. Address requests to The Institute of Radio Engineers, 1 East 79th Street, N.Y. 21, N.Y.

Notice to Authors: Address all papers and editorial correspondence to H. E. Tompkins, The Moore School of Electrical Engineering, 200 South 33rd Street, Philadelphia 4, Pa. To avoid delay and inconvenience in the processing of manuscripts, please follow the procedure suggested on the back cover, under Information for Authors.

COPYRIGHT © 1959—THE INSTITUTE OF RADIO ENGINEERS, INC.
Printed in U.S.A.

All rights, including translation, are reserved by the IRE. Requests for republication privileges should be addressed to the Institute of Radio Engineers.

Richard O. Endres

CHAIRMAN, 1959–1960

At the PGEC Administrative Committee Meeting in March, Richard O. Endres was elected Chairman for the year 1959–1960. Mr. Endres graduated from Purdue University in 1948 and joined the Engineering Products Department of the Radio Corporation of America. He did early work on point contact transistor circuits, and at RCA Laboratories in Princeton did development work on discrete type electrostatic storage systems. He was responsible for development and construction of the magnetic core memory system for the RCA BIZMAC computer, and was then appointed manager of the computer circuits group investigating high-speed transistor logic circuits and advanced magnetic core storage techniques.

In 1955 he joined Rese Engineering Company to direct its engineering program as chief engineer. He was elected president of the corporation in 1958 following the company's entry into the digital equipment field, and presently occupies that position.

He is co-author of "Transistor Electronics," published by Prentice-Hall in 1955, and is author of a number of other transistor circuits papers. He was active in the Philadelphia Chapter of the PGEC, serving as Program Chairman, and participated in the Philadelphia Chapter's Symposium on Computer Fundamentals. He has served on the PGEC National Administrative Committee for the past two years and was Vice-Chairman of PGEC during the year just ending.

As the new Editor, I extend the new Chairman congratulations, and express the hope and expectation that for all of us in PGEC, 1959–1960 will be a rewarding and prosperous year, under Dick Endres' guidance.

—H. E. TOMPKINS

The Chairman's Column

To Members of the PGEC

SINCE my last report to you there has been another Professional Group meeting in New York and two more meetings of the PGEC Administrative Committee. First, the PG meeting in New York. As with the last several such meetings, this one was again attended for me by Charles Rosenthal (Administrative Committee member from New York). The following summary is based on his report to me.

The Professional Group on Communications Systems and the Professional Group on Vehicular Communications have mutually decided that they do not wish to merge. Arrangements have been completed to distribute the WESCON Convention Record in the same manner as the IRE National Convention Record. There will be no free automatic distribution to group members, but group members who wish to purchase some or all of the WESCON Convention Record may do so at a substantially reduced price. The proposed new Professional Group on Space Electronics has been abandoned in favor of incorporating this area into the charter of the Professional Group on Telemetry and Remote Control, which has changed its name to the Professional Group on Space Electronics and Telemetry.

The IRE policy with respect to membership in international bodies was reiterated; this is of particular interest to the PGEC because of its negotiations (through I. L. Auerbach and the NJCC) in behalf of an International Federation of Information Processing Societies. There appears to be no difficulty in having the PGEC represent the name of the IRE in such an international organization. The Professional Group on Automatic Control already represents the IRE to the International Federation on Automatic Control; this representation is through an intermediate organization in this country known as the American Automatic Control Council. The Professional Group on Medical Electronics is also negotiating in behalf of an international federation in its field of interest.

A meeting of the Administrative Committee occurred during the 1959 Western Joint Computer Conference at San Francisco, the first week of March. Since only a few members were present, no business was transacted. Some of the problems of the PGEC were informally discussed. An interesting sidelight of the meeting was that the Chairman had received so many proxy votes that he lacked only one vote of being a quorum all by himself.

The Annual Meeting of the Administrative Committee was held during the National Convention, the third week of March. The most important events at this meeting are the annual report of each committee, and the election of new officers and administrative committee members. Vice-Chairman Dick Endres has ascended to the Chair, and Dr. Arnold Cohen of St. Paul is the newly elected Vice-Chairman. Five new members of the Administrative Committee were elected. They are:

Baltimore—Dr. L. G. F. Jones
Detroit—Dr. E. Calvin Johnson
Binghamton—Dr. Yates M. Hill
New York—Matthew J. Relis
San Francisco—R. W. Melville.

The retiring members of the Administrative Committee are:

New York—David C. Bomberger
Poughkeepsie—Dr. Werner Buchholz
Washington—J. Claude LaPointe
Chicago—Henry P. Messinger
Dearborn—Robert A. Roggenbuck.

We of the PGEC are certainly grateful to each of these men for his several years of service to the PGEC, and we welcome them to the ranks of the elder statesmen of the Group. On such experienced people the PGEC can depend for advice and guidance, and from them also draw an occasional chairman for a necessary ad hoc committee.

Highlights of the committee reports: abstracts are coming along well with the second set appearing in this issue; write to Committeeman Frank Heart and let him know what your reaction is to this effort . . . the bibliographic effort is getting onto a more formal basis with meetings proposed for the several interested organizations . . . membership is somewhat over 7000 and the treasury balance at year end was just under $30,000; we expect this balance to drop, however, during the coming year when the new IRE financing policy comes into effect . . . the Constitution and Bylaws Committee still has a few bugs in its revised drafts—it is expected that the new draft will be ready for approval by summer . . . the Awards Committee (conducted by the old and the new Vice-Chairmen) have or are submitting recommendations for the various IRE prizes and awards . . . the editor reviewed his plans for improving the Transactions and getting publication back on schedule.

There was limited discussion of a major problem which the PGEC currently faces. In order for future international meetings to be held, it is necessary to organize a suitable international federation catering to computing societies of the various countries. Such a federation will only have one representative from each country and, therefore, some mechanism must be found by which the many computer and computer-oriented societies of the U. S. can be represented by one delegate. This is the reason for the existence of the American Automatic Control Council. One possibility for such a

"society of U. S. Computer Societies" is an enlarged National Joint Computer Committee. There is also the possibility of forming a new organization for this purpose. The PGEC also needs to review its domestic position; with other PG's critically reviewing their areas of interest in view of technological advances and the opening of new fields, it is time for the PGEC to introspect and determine which position it wants to have in the U. S. computer fraternity and to move in that direction. It might even be desirable to consider modifying the name of the Group.

The two new PGEC delegates to the NJCC are Dr. Werner Buchholz (now senior past chairman of the PGEC), and Dr. Willis H. Ware (junior past chairman by the time this column is in print). Their terms will run through March, 1961. The other two PGEC delegates to NJCC are R. D. Elbourn, and Harry H. Goode; Chairman-elect Endres is an *ex-officio* member.

I would like here to express my appreciation to all members of the Administrative Committee and the Chairmen of the various standing committees for their fine cooperation during my term as Chairman. I want also to acknowledge the wonderful job which Secretary-Treasurer Bill Speer has done over the past year. And it certainly goes without saying that except for my own secretary, Mrs. Dorothy Crabb, and Bill's secretary, Claire Burks, things could not have gone so smoothly and efficiently for us.

WILLIS H. WARE
Chairman 1958–1959

The 1959 Solid-State Circuits Conference, held February 12 and 13, in Philadelphia, Pa., on the campus of the University of Pennsylvania, included many papers of particular interest to PGEC members. We present four of them on the next several pages, and expect to include others in the September, 1959, issue. —*The Editor*

Thin-Film Memories*

ERIC E. BITTMANN†

Summary—**A small random-access memory using deposited magnetic thin films as storage elements, and with a cycle time of one microsecond, is described. Information is read from or written into the memory by linear or word selection techniques. The addressing, driving and sensing circuits are transistorized. The deposited thin films are 2000 Å thick, switch in 0.1 μsec and generate a 5-mv output signal in the sense winding. A sense signal is obtained of opposite polarity from a selected element when a "1" or a "0" is read out. A memory-plane wiring configuration has been selected which is least susceptible to noise.**

Introduction

INVESTIGATION of the flux-reversal mechanism in ferromagnetic thin films, when less than 5000 Å thick, has revealed a predominant spin-rotational remagnetization rather than a domain-wall movement as observed in ferrite cores.[1–5] The films show a preferred direction of magnetization as demonstrated by characteristic hysteresis loops in various directions.[2,6] The switching phenomena has been observed also by fixed pickup and driving loops, where the sample was rotatable.[7] These measurements reveal a predominant spin-rotational switching which can be interpreted by a magnetic dipole representation. These films are bistable magnetic elements[8] and can be produced by vacuum deposition.[9] They can replace the small ferrite cores used in memories for digital computers.[10] The thin-film elements in the described memory are deposited Ni-Fe films, 2000 Å thick. The deposition is performed in vacuum onto hot glass substrates under the influence of a magnetic field. The deposition of the Ni-Fe alloy through a metal mask produces memory planes with round spots 3/16-inch diameter. The film's preferred direction is determined to a large extent by the orientation of the magnetic field during deposition.[11] In a 60-cycle loop tracer the films show a square hysteresis loop along the preferred direction, and 90° to it show an almost linear hysteresis loop.

Behavior of Thin Film

For simplification, the magnetization of each deposited area can be represented by a compass needle or a magnetic dipole (Fig. 1). In the absence of any outside magnetic field, the needle has two stable states. Both are parallel to the preferred magnetic direction. The effect of the earth's magnetic field is secondary, and can be eliminated by a mumetal shield. The two states of the magnetic dipole are called the N state and P state. If a current of sufficient magnitude is sent through a conductor lying on top of the film, its magnetic field causes the dipole to rotate out of the P state or N state toward alignment with the drive field. When the current is removed, the drive field disappears and the dipole finds itself in an unstable position. It will rotate toward the closest stable state. In the case shown in Fig. 1, this would be the P state. For use as a memory element the P state could be defined as the storage of a "1" and the N state as the storage of a "0"

Memory Array

The sense conductor is always at right angles to the drive conductor to minimize noise pickup from the drive-current pulses.

* Manuscript received by the PGEC, February 13, 1959; revised manuscript received April 16, 1959. This paper was presented at the 1959 Solid-State Circuits Conference, Philadelphia, Pa., February 12–13, 1959.

† Research Center, Burroughs Corporation, Paoli, Pa.

[1] C. Kittel, "Theory of the structure of ferromagnetic domains in films and small particles," *Phys. Rev.*, vol. 70, pp. 965–971; December, 1946.

[2] C. D. Olson and A. V. Pohm, "Flux reversal in Ni-Fe films," *J. Appl. Phys.*, vol. 29, pp. 274–282; March, 1958.

[3] F. M. Humphrey, "Transverse flux change in soft ferromagnetics," *J. Appl. Phys.*, vol. 29, pp. 284–285; March, 1958.

[4] D. O. Smith, "Static and dynamic behavior of thin permalloy films," *J. Appl. Phys.*, vol. 29, pp. 264–273; March, 1958.

[5] F. M. Humphrey and E. M. Gyorgy, "Flux reversal in soft ferromagnetics," to be published.

[6] R. L. Conger, "Magnetization reversal in thin films," *Phys. Rev.*, vol. 98, pp. 1752–1754; June, 1955.

[7] J. D. Blades, "Thin film magnetization," unpublished report.

[8] R. G. Alexander, "Anisotropy measurements in Ni-Fe thin films," *J. Appl. Phys.*, vol. 30, pp. 266S–267S; April, 1959.

[9] M. S. Blois, Jr., "Preparation of thin magnetic films and their properties," *J. Appl. Phys.*, vol. 26, pp. 975–980; August, 1955.

[10] A. V. Pohm and S. M. Rubens, "A compact coincident current memory," *1956 Proc. EJCC.*

[11] J. D. Blades, "Stress anisotropy in Ni-Fe thin films," *J. Appl. Phys.*, vol. 30, pp. 260S–261S; April, 1959.

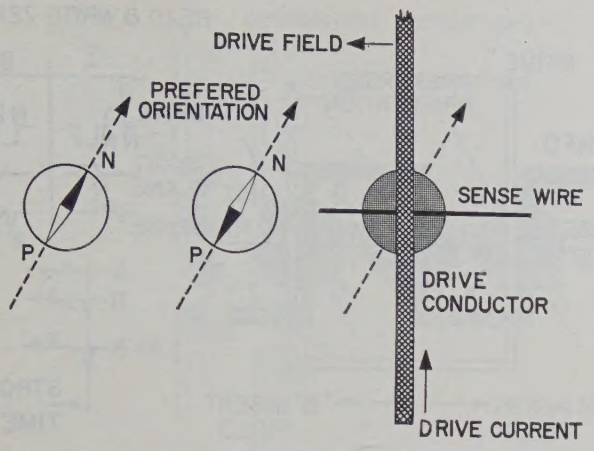


Fig. 1—Magnetic dipole representation of a thin-film element.

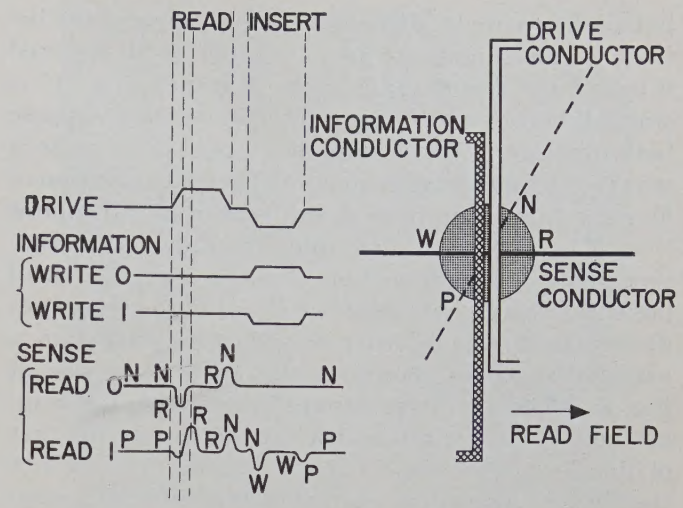
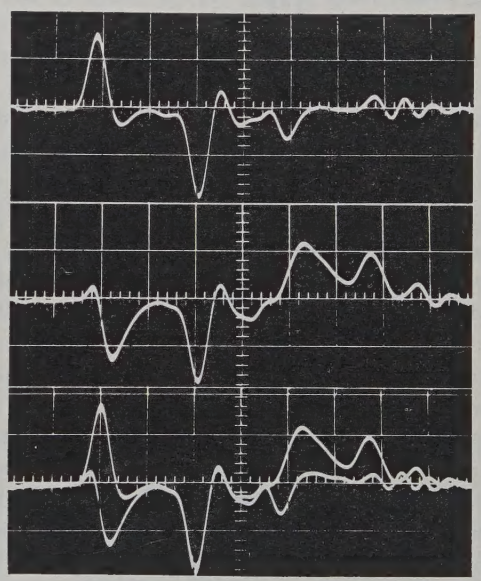


Fig. 2—Thin-film memory bit; the three associated conductors, drive, information, and sense; current waveforms and readout signals.



Sense Signal "1's" and "0's"
0.1 μsec/cm 5 mv/cm

Fig. 3—Sense signal of a thin-film spot.

The drive and information conductors can be arranged in two ways. One, not used in the described memory, has the information and sense conductors parallel to each other and at right angles to the drive conductor. The preferred direction should then be parallel with the drive conductor. This was tried and found workable, but the selection ratio of the film was inferior to the selection ratio which is obtained in the second method, where the films are driven by two parallel drive fields. Furthermore, larger variations in the film's characteristics are tolerable in the second method. Fig. 2 shows the drive and information conductors to be parallel to each other and at right angles to the sense conductor. The drive and information conductors are shown next to each other; actually one lies on top of the other.

The state of a thin-film element is sensed by subjecting it to a magnetic field which tends to rotate it toward the R state. This field is called the read field and can be as large as desired. The rotation of the dipole from N to R induces a negative signal in the sense conductor and a rotation from P to R induces a small negative signal followed by a larger positive signal. After termination of the read field the dipole will fall to the N state, inducing another positive signal in the sense winding. During the write operation two magnetic fields are applied, an insert field on the drive conductor and a write "1" or write "0" field on the information conductor. The write "0" field subtracts from the insert field and the write "1" field adds to the insert field. The force of the insert field is $\frac{2}{3}$ the magnetic threshold and the write "1" or "0" is $\frac{1}{3}$ of the threshold. When writing a "0" the dipole remains in the N state during the write operation, due to the application of only $\frac{1}{3}$ of the threshold field. When writing a "1" the dipole rotates to W and after the fields are removed falls to the P state. The signals induced in the sense winding are not of interest, but are both of negative polarity. If the two possible signals obtained during a read operation are examined, a difference in time between the peak signal of a sensed "1" and a sensed "0" of about 50 mμsec. is observed. Fig. 3 shows actual oscilloscope pictures of "1" and "0" readout signals and the time difference can be observed. The overshoot could be interpreted as a "zero" if somewhat large. The overshoot is only to be expected when the film used behaves exactly as a magnetic dipole. The film's rotational switching induces a sense signal at the beginning and at the end of the read current pulse and for short pulses these signals will be close together and interpretation of a "1" or "0" might become difficult.

If storage of a bit is performed in two spots instead of in one spot (Fig. 4), and if these spots are not adjacent,

but as shown are in alternate locations, some of the before-mentioned problems are eliminated. A "0" is stored if both films A and B are in the N state and a "1" is stored if both A and B are in the P state. Now suppose both films are in the N state and a read drive pulse is sent through the drive conductor. The magnetization of film A rotates from N to R, but that of film B rotates from N to W. The rotation from N to R induces a negative signal and rotation from N to W induces a small positive followed by a negative signal in the sense conductor. Both signals from A and B combine into a negative signal as shown on the lower right side of Fig. 4. When the drive current is removed, the magnetization of film A rotates back to the N state and that of film B rotates to the P state. To write a "0" back into the films, an insert drive current of $\frac{2}{3}$ full switching magnitude is sent through the drive conductor and at the same time a write current of $\frac{1}{3}$ switching magnitude is sent through the information winding which subtracts from the insert field in film A, but adds to the field in film B. Therefore, film A remains in the N state, but film B rotates from P to R and upon removal of the field falls to N and the write cycle is completed.

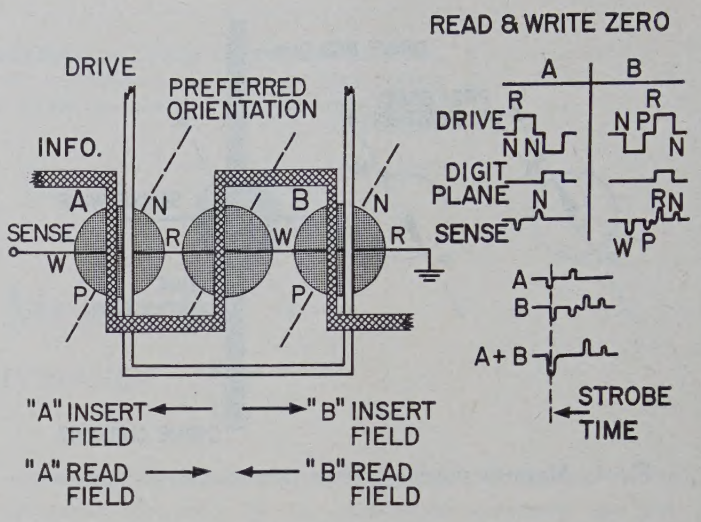


Fig. 4—Storage of a bit in two spots; read and write "0".

If both films are in the P state (Fig. 5), and a read current is applied, the rotation from P to R in film A induces a small negative followed by a positive signal and rotation from P to W in film B induces a positive signal and the combination of the two produce a sense signal exactly opposite in polarity, but of the same duration and shape as the read "0" signal. Upon removal of the drive field, film A falls back to N and film B falls to P again. To write the "1" back into the film the insert current and a write "1" current of opposite polarity are sent through their conductors. The magnetic fields which add in film A rotate it to W, and those which subtract in film B leave it in the P state. When the fields are removed film A falls to P and film B remains in P. The cycle is completed and the "1" is written into the films. The wiring arrangement of a single plane is shown in Fig. 6. Linear, or word selection drive is used, all films of the same address being driven at the same time by a full switching current. All wiring is such that the air inductance is kept to a minimum, and wiring loops are kept as small as possible.

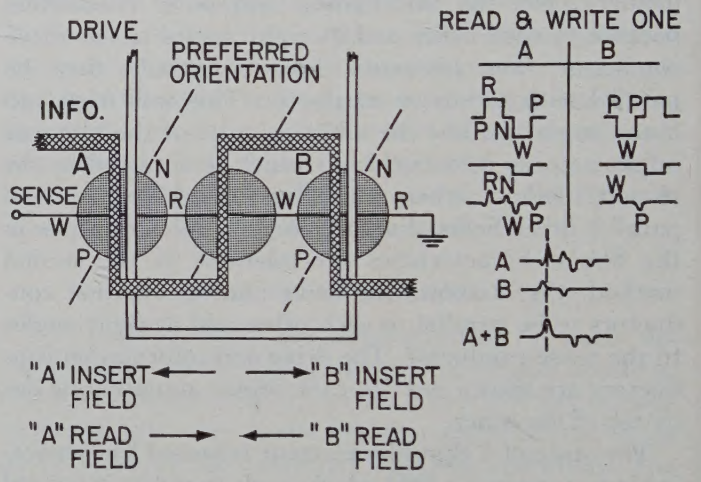


Fig. 5—Storage of a bit in two spots; read and write "1".

A single plane as shown cannot be used in the manner described. There is a noise problem present. The sense conductor is at right angles to the drive conductor and noise cancellation between these two conductors is satisfactory. The third conductor, the information winding, is added to the plane. It is in part parallel to the sense winding and part parallel to the drive winding. A drive current induces a noise current in the information conductor, and this noise current flowing in that winding, in turn induces noise in the sense winding. A second memory plane is needed, in which the noise current flowing in the information conductor reverses its direction with respect to the sense conductor and this is accomplished by an arrangement shown in Fig. 7. The orientation of the information conductor is reversed on the second plane. This wiring arrangement has three nice features. First, it minimizes the noise; second, it insures that all outputs are positive for the reading of a stored "1" and negative for a stored "0" (provided that the polarity of the information current is alternated between the storage of a "1" in an even address and storage of a "1" in an odd address); and third, the noise induced during the write cycle by the information write current is self-cancelling so that no large pulses appear in the sense winding. This feature is of importance when it is desired to squeeze the cycle time below 1 μsec. No ringing noise or storage effect are carried over into the next read cycle. Thus for good signal-to-noise ratios, the memory planes should be in pairs.

Memory Fabrication

Sample memory planes were built, having 4 by 7 depositions of 3/16-inch diameter spots on $\frac{1}{4}$-inch

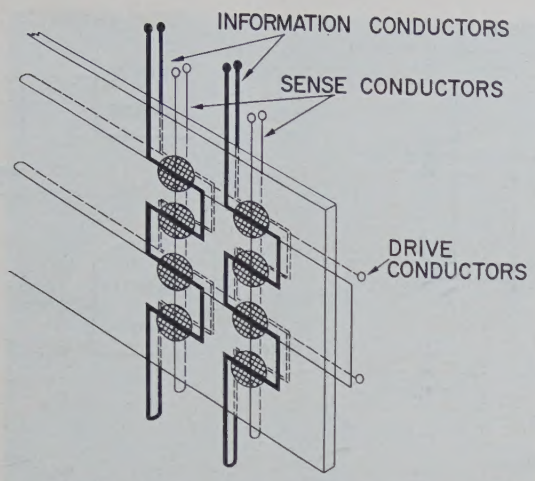


Fig. 6—Wiring arrangement of a thin-film memory plane.

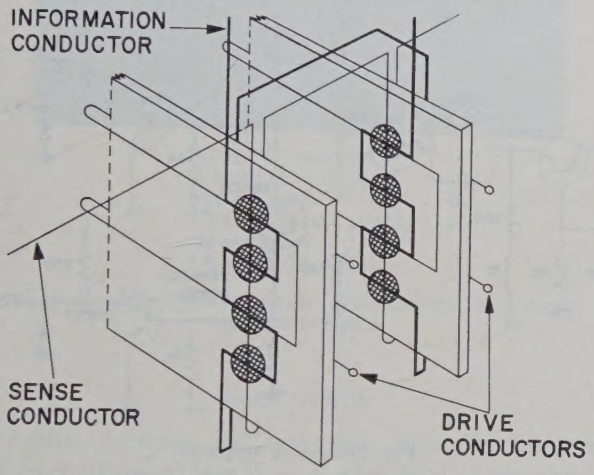


Fig 7—Interwiring of two thin-film memory planes.

centers along the narrow edge, and on $\frac{3}{8}$-inch centers along the long edge of the glass. The planes were assembled in the following way. (Fig. 8.) The seven information conductors were plated first onto a glass epoxy board. Conductors were 1/16-inch wide. These conductors were then coated by a thin layer of insulation by two or three coats of glyptol or by a layer of thin mylar. The four 1/16-inch wide drive conductors were then plated on top and if mylar insulation was used, 1/16-inch wide brass strips were cemented on. Two identical epoxy boards were made. The thin film slide was mounted into a frame, and seven sense windings were wound on. Number 36 magnet wire was used for the sense conductor. The wire rather than a plated conductor was used to minimize capacity between the sense and drive conductors to keep noise signals low. The prewired memory plane was then sandwiched between the two epoxy boards (Fig. 9), the sense wires were soldered to the provided terminals and the memory plane was complete. Two planes were wired together by straight bus bars and then tested. No adjustment or addition of balancing loops was necessary. Fig. 10 shows two 4 by 7 planes interwired.

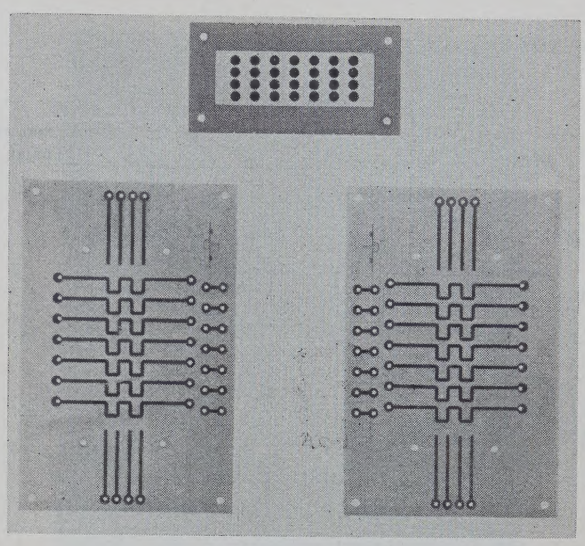


Fig. 8—Memory plane parts before assembly.

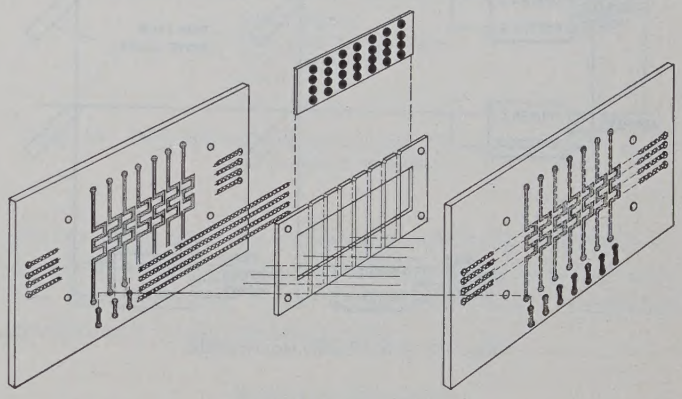


Fig. 9—Expanded view of a thin-film memory plane.

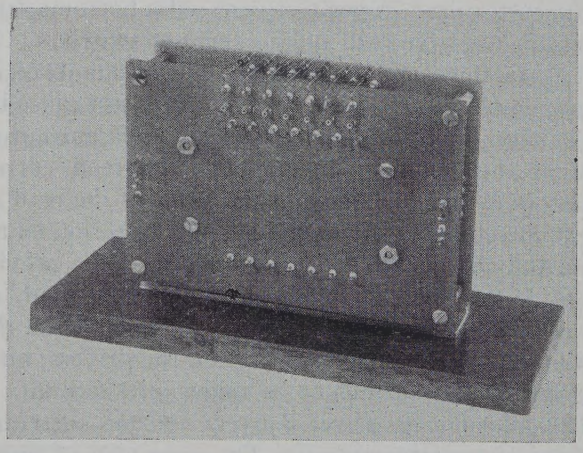


Fig. 10—Two 4 by 7 memory planes interwired.

Circuitry

The peripheral circuitry is completely transistorized, the current drivers delivering 1-amp pulses with rise time of 0.15 μsec. Three *NPN* type 2N576 transistors in parallel, with a current rating of 400 ma each are used in the output stage. As a current switch, three *PNP* type 2N580 transistors, with a current rating of 400 ma

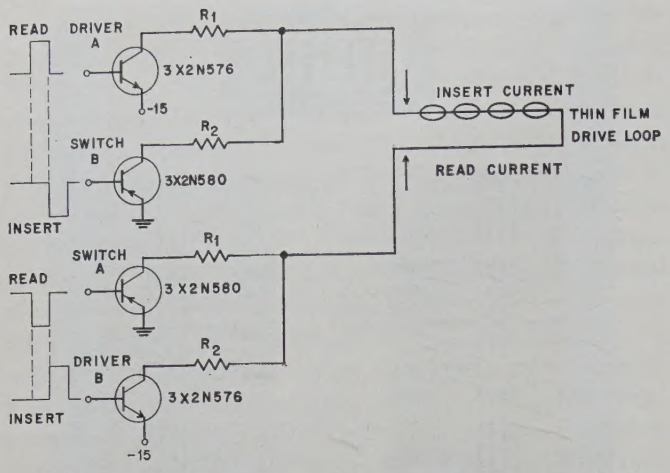


Fig. 11—Read-insert current drivers.

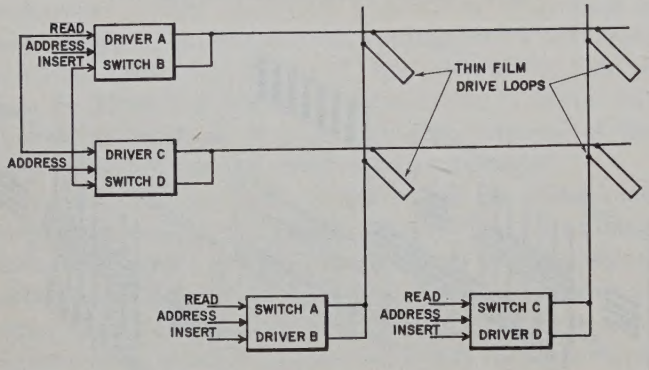


Fig. 12—2 by 2 selection matrix.

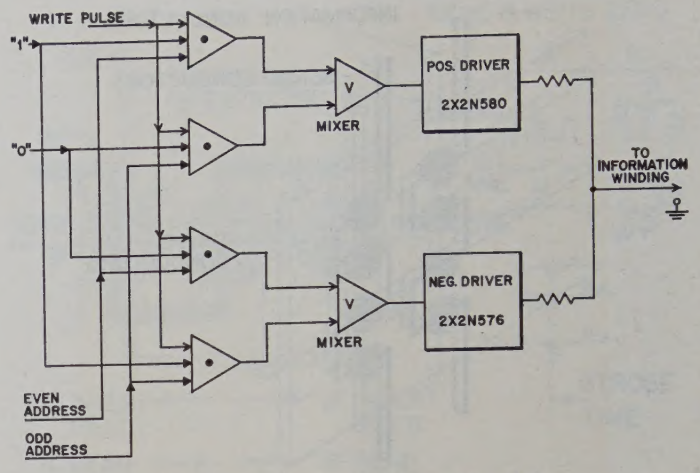


Fig. 13—Logic diagram of information driver.

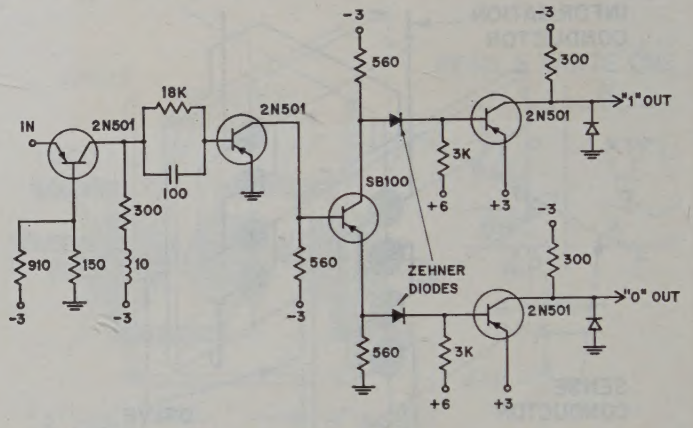


Fig. 14—Sense amplifier.

each are connected in parallel. Fig. 11 shows the driving arrangement used. A positive read pulse bottoms driver A, and a negative read pulse energizes switch A. Current flows through the thin-film loop in one direction. After completion of the read pulse the insert pulse bottoms driver B and switch B, and the insert current will flow in the opposite direction to the read current. Resistors R_1 and R_2 determine the value of the read and insert currents. Fig. 12 shows a 2 by 2 selection matrix, using the current driver and switches of the previous figure. In the sample, memory diodes were used, and the selection scheme operated satisfactorily. The thin-film load imposes little back emf on the drivers; air inductance is all that has to be taken into account.

The information driver delivers 400 ma currents of positive or negative polarity. The same type of transistors as in the current drivers were used and Fig. 13 shows the block diagram, indicating the gating required to insure the writing of the correct information into any selected address.

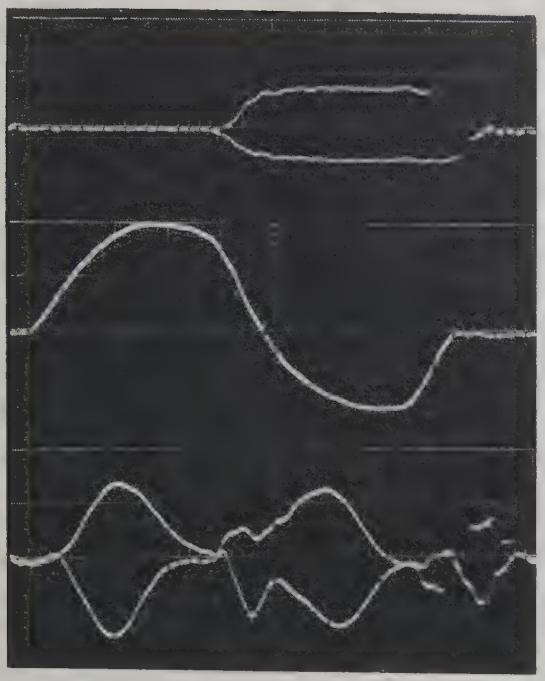
The sense amplifier, Fig. 14, uses 5 transistors and amplifies the 5 mv input signal to a 3-volt level. The common-base input stage matches the low source impedance of the sense winding. The Zener diodes shift dc levels of the output signal to the desired 0 to +3 volt levels.

Fig. 15 shows a drive current pulse, the write "1" and "0" current pulses and the resulting output signals. Fig. 16 shows the sense signal for the reading and writing of "1" and "0".

Memory Operation and Feasibility

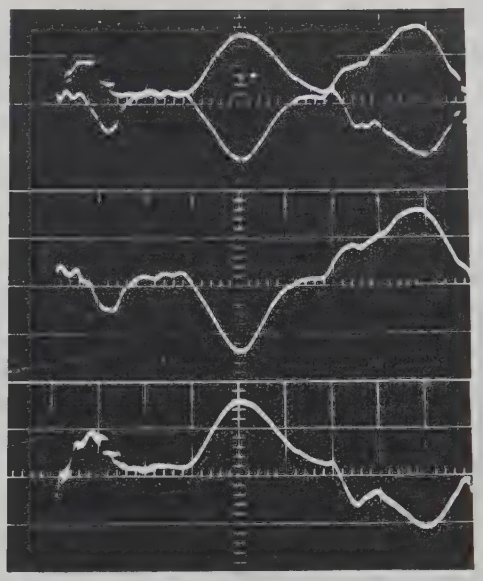
The switching time of the thin film is very short. All films switch during the current rise time and the limitations on speed are in the drivers. In this memory model, current pulses are one-third μsec wide for the read and insert pulses, which leaves one-third μsec time for the decoding. A cycle time of 1 μsec is possible without any crowding.

The films are made by vacuum deposition. Several memory planes are made at one evaporation process, and quality and uniformity of the films are good. Uniformity of different runs is within 20 per cent. Usually, if one spot on a plate is found to be good, all on that plate are good and, furthermore, all plates made in that run are good. It was due to this quality and uniformity that feasibility of a thin film memory could be proven.



Write "1" and "0"
Current 0.5 amp/cm
Drive Current
0.5 amp/cm 0.1 μsec/cm
Sense Signal
5 mv/cm 0.1 μsec/cm

Fig. 15—Information current, drive current, and sense signal waveforms.



Sense Signal
5 mv/cm 0.1 μsec/cm
"1's" and "0's"
"0's"
"1's"

Fig. 16—Thin-film readout, sense signal (2 spots). (Desired signal in center of sweep.)

Conclusion

Results obtained indicate that thin film memories with cycle times of less than a microsecond are possible. The films show better physical properties than those of ferrites. They operate over a wide temperature range (−40° to 140°F) and switch in the millimicrosecond range. The fabrication of thin-film memory planes seems to be simple. The films, the insulators, and the conductors will eventually all be deposited. Completed memory planes fabricated by deposition methods will then be possible, thus reducing the memory cost. This fact plus the fast operating speed make thin films a desirable component in digital computers.

Acknowledgment

The contributions made to this project by the Ferromagnetic Thin Film Group under Dr. John D. Blades are gratefully acknowledged. Discussions with Dr. F. B. Humphrey of Bell Laboratories, and with Dr. D. O. Smith and J. Raffel, both of The Massachusetts Institute of Technology, have been helpful.

Integrated Devices Using Direct-Coupled Unipolar Transistor Logic*

J. T. WALLMARK† AND S. M. MARCUS‡

Summary—This article presents material that is new in three areas. First, a new logic system using directly-coupled unipolar transistors is analyzed. It is shown that unipolar transistors have important advantages over bipolar transistors in speed, tolerance of stray signals and noise, and device miniaturization. Second, devices of extreme miniaturization built by an integrated device design and using this logic system are described. Third, how the passive components of the system, in this case resistors, have been integrated into the semiconductor devices is described.

I. Introduction

IT has been suggested that perhaps the greatest importance of the transistor lies in the possibility of a more efficient replacing, or complementing, of mental processes as is done in computers and automation equipment. While in this application the small dimensions of the transistor, as well as its low power requirements, constitute major steps forward compared to electron tubes, the ultimate potential of the transistor in this respect has not yet been determined and certainly not yet been reached. One approach to this ultimate potential of the transistor is the integrated devices concept which has been described elsewhere.[1] The requirement in this approach that the transistor be made in a form which allows a very high space packing factor can possibly be best met by the use of integrated devices utilizing unipolar transistors, as outlined in this paper.

Briefly described herein is a computer logic system called the direct-coupled unipolar transistor logic (DCUTL). While it is the chief advantage of DCUTL that large arrays and matrices in extremely compact form are very promising (and this aspect is stressed), DCUTL may as well be used with individual units or combinations of a few individual units. Even in this case it is believed that DCUTL offers several advantages over existing logic systems.

It is shown that unipolar transistors may be made in large arrays, where one array may constitute a complex logic circuit. The suggested fabrication technique is the lapping-diffusion process developed by Nelson.[2] A technique for wiring and encapsulation of such complex units is described.

* Manuscript received by the PGEC, February 26, 1959; revised manuscript received April 13, 1959. This paper was presented at the 1959 Solid-State Circuits Conference, Philadelphia, Pa., February 12–13, 1959.

† RCA Laboratories, Princeton, N. J.
‡ RCA Defense Electronic Products Div., Camden, N. J.

[1] J. T. Wallmark, "Integrated Electronic Devices—A New Approach to Microminiaturization," "Aviation Age, Handbook on Research and Development," p. F6; 1958–1959.

[2] H. Nelson, "The preparation of semiconductor devices by Lapping and diffusion techniques," PROC. IRE, vol. 46, pp. 1062–1067; June, 1958.

II. Basic Device Factors

A. The Unipolar Transistor

Extensive treatments of the unipolar transistor[3,4] are available and only a brief description will be pertinent here. Fig. 1 shows a schematic picture of a unipolar transistor with common nomenclature indicated. The resistance of the channel from the source to the drain may be varied by varying the gate voltage. When the gate is made negative with respect to the channel (source and/or drain) the junction depletion layer grows and encroaches upon the channel making it narrower. Therefore the resistance of the channel increases. When the gate voltage is made less negative, the channel resistance decreases. The change in resistance for the dimensions given may be obtained from Fig. 2 which gives some typical characteristics of a unipolar transistor. The gate is characterized by a saturation current of much less than 1 μa with the dimensions in Fig. 1, and a capacitance which is shown in Fig. 3 as a function of gate voltage.

B. The Basic Unit of DCUTL

Fig. 4 shows the basic logical unit of DCUTL consisting of one unipolar transistor and one resistor. The resistor leading to a negative potential insures that the transistor is held in the OFF (high resistance in the channel) position. A positive pulse on the gate bringing the potential to a value only slightly negative corresponds to the ON (low resistance in the channel) position. This operation is analogous to a voltage operated relay and, as shown later, ordinary relay logic will be used in building complex networks.

The unipolar transistor can be made, not only in the p^+-n version described above, but also in an n^+-p version. For distinction this version will be drawn filled-in in the schematics following.

In logical circuits it is also necessary to have the inverse of the function of Fig. 4. This is obtained by reversing the current in the output (compare Fig. 6, which illustrates a chain of regular units, with Fig. 7, which shows a chain of inverter units).

C. The Equivalent Circuit

With the dimensions shown in Fig. 1 and with silicon the negative voltage necessary for OFF may be approxi-

[3] G. C. Dacey, and I. M. Ross, "Unipolar field-effect transistor," PROC. IRE, vol. 41, pp. 970–979; August, 1953.

[4] W. Shockley, "Unipolar field-effect transistor," PROC. IRE, vol. 40, pp. 1365–1376; November, 1952.

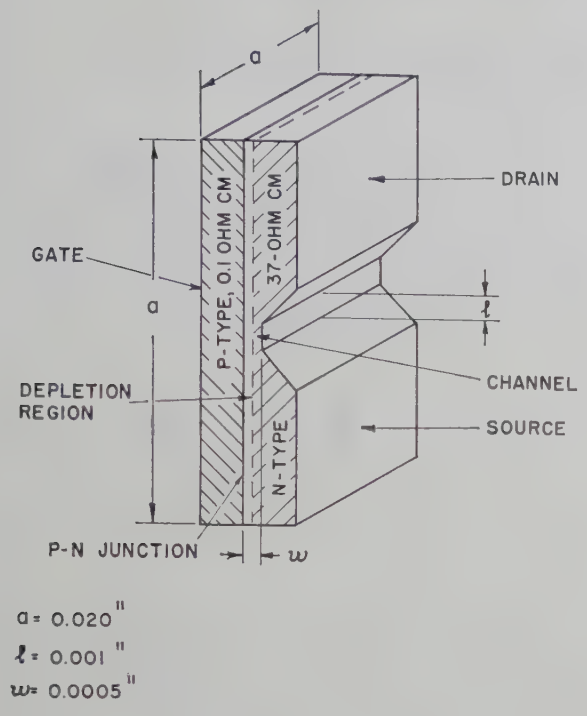


Fig. 1—Unipolar transistor.

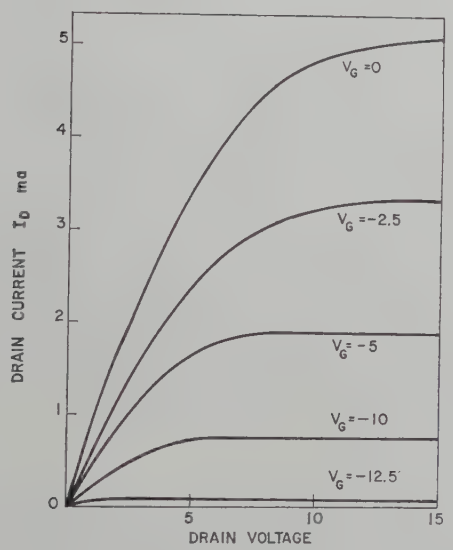


Fig. 2—Typical characteristics of unipolar transistor.

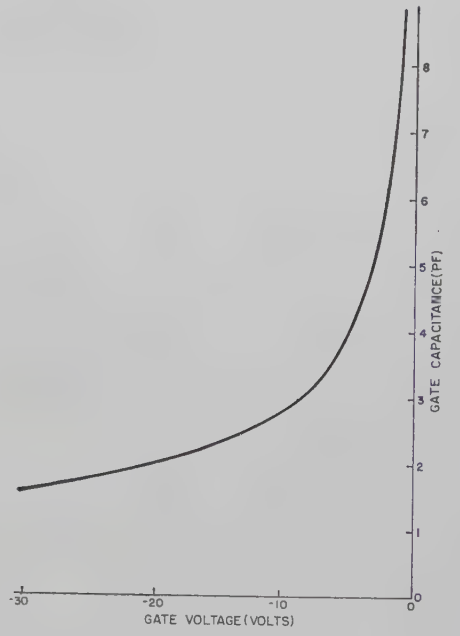


Fig. 3—Gate capacitance as a function of gate voltage.

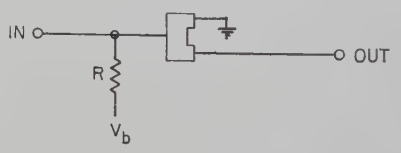


Fig. 4—Basic unit of DCUTL.

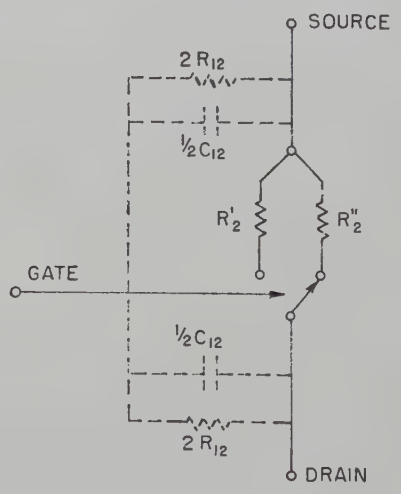


Fig. 5—Equivalent circuit of basic unit.

mately 15 volts. Assuming, as will be elaborated later, that the gate voltage for OFF is −15 volts and for ON −5 volts, the equivalent circuit becomes what is shown in Fig. 5, with the following values

$$R_2' = 2000 \text{ ohm}$$

$$R_2'' \geqq 20{,}000 \text{ ohm}$$

$$R_{12} = 10 \text{ megohm}$$

$$C_{12} = 4\ \mu\mu\text{F}.$$

The coupling elements from input to output, the junction capacitance C_{12} and the reversed-biased junction resistance R_{12} (shown dashed in Fig. 5) will be neglected, assuming low frequency operation and a gate voltage that is always negative. Their influence will be evaluated in Section *J*.

D. Choice of Material and Dimensions

The requirements on the unipolar transistor element from different points of view such as speed of operation, miniaturization, ease of fabrication, etc., are largely contradictory. While speed of operation favors germanium of relatively low resistivity, miniaturization

favors silicon with relatively high resistivity, and ease of fabrication excludes low resistivity altogether. As the main goal of the work reported here is extreme miniaturization, silicon was chosen for two reasons. First the power developed per unit for given voltage and and channel dimensions is proportional to the product $\epsilon\mu$ and therefore 3.9 to 5.3 times smaller for silicon than for germanium. Second for given dimensions, silicon tolerates about three times higher temperature above the ambient and therefore three times higher heat flow. Further contributing to this is the fact that thermal conductivity of silicon is more than two times that of germanium. Silicon also allows higher resistivity at room temperature and therefore may allow a future reduction in power per stage.

The next important consideration is that of ease in fabrication. With the present state of fabrication techniques it is not considered desirable to maintain a channel width less than 0.0005 inch. At the same time a pulse height of approximately 10 volts was required by considerations of compatibility with existing transistorized equipment. This leads to a minimum resistivity of 37 ohm cm.

With this resistivity and the dimensions of Fig. 1 the ON resistance is approximately 2000 ohms. The OFF resistance can be made very high by choosing sufficiently high negative bias, close to or past pinch-off. We will assume a pinch-off voltage of -15 volts, a load resistance of 4000 ohms, and a battery voltage of 15 volts. Then the gate bias values become -5 volts for ON ($+5$ volts for a unit with complementary symmetry), and -15 volts for OFF ($+15$).

For many applications it may not be necessary to drive to pinch-off. When a unit is turned off most of the output signal has been obtained when the value of the transistor output resistance has reached a few times that of the load resistance, and further increase in the output resistance may not be essential.

E. Series and Parallel Connection

Fig. 6 shows a row of series-connected basic elements. As is shown, positive and negative supply voltages alternate for alternate stages in order to allow a reverse bias to remain across the junction in the ON state. Consequently also the p^+-n and the n^+-p versions alternate to keep the ON or OFF state consistent through the chain. The number of basic units that can be connected in series is unlimited as long as the voltage level for ON or OFF is convergent through the series, *i.e.*, in spite of a deviation in one stage the level returns to its correct value in subsequent stages. This will be treated later in the discussion of stability, Sections *G* and *H*.

Another alternative for series connection is shown in Fig. 7. Here ON and OFF elements alternate along the chain, each element behaving as an inverter.

The number of units that can be connected in parallel is limited only by impedance considerations, *i.e.*, the combined impedance should allow signal stability as described in Sections *G* and *H*. This limit is very large, 100 or more, when gates are paralleled. When outputs are paralleled as in Fig. 12 the limit is also large, >10, assuming that all the elements are driven to pinch-off. When the outputs are series-connected, as in Fig. 13, the load resistance is increased in proportion to the number of units. When this is done the limit here is also >10.

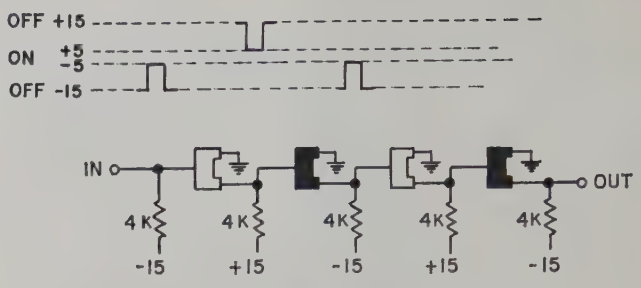


Fig. 6—Series connection.

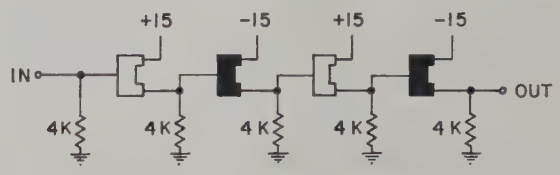


Fig. 7—Series connection of inverters.

When a large number of units are connected, either in series or in parallel, the speed of response of the network will of course decrease over that of an individual unit. This will be considered in Section *I*.

F. Output Signal

The output signal from a DCUTL network which would be available for driving further circuits, such as display panels, etc., is of course comparable in magnitude to the input signal. With the dimensions chosen this would amount to 10 volts across 4000 ohms.

G. DC-Level Stability

One of the requirements of a logic system is that the signal does not deteriorate appreciably on its way through the system. Otherwise repeated amplification and reshaping of the signal has to be introduced. This is costly. Unipolar transistors, which themselves are capable of amplification, provide internal restoring of the signal, as will now be shown.

Let us first consider dc-level stability, *i.e.*, whether the two states ON and OFF, represented by two negative gate potentials, V_{ON} and V_{OFF}, will propagate through a chain of units such as shown in Fig. 6 without deterioration. This requires the next stage to show a smaller deviation, if the signal level at any point in the chain deviates from its proper value; the subsequent stage a still smaller one, etc. This condition may be called "self-locking."

It is shown in Appendix I that the ON state with a small negative gate voltage V_{ON} is indeed self-locking. The OFF state with a large negative gate voltage is not

self-locking in the first order theory. However, by the use of overdrive, *i.e.*, using a sufficiently large OFF voltage to insure that the gate is driven past pinch-off the OFF state may be made self-locking.

H. AC-Level Stability

Each unipolar transistor constitutes an amplifier, and with many of them cascaded, noise from the input may be amplified through the chain and may ultimately reach the signal level. In the OFF state this is not a problem as the amplification is negligible. In the ON state the condition that the amplification in each stage be equal to or less than unity gives (see Appendix II),

$$\frac{g_m}{\frac{1}{R} + g_m \frac{\sqrt{V_P} + \sqrt{V_D - V_G}}{\sqrt{V_D - V_G} - \sqrt{-V_G}}} \leqq 1,$$

where

g_m = transconductance of the unipolar transistor
V_G = gate voltage
V_D = drain voltage
V_P = pinch-off voltage
R = gate load resistance.

This condition is fulfilled by transistors with the dimensions chosen.

I. High-Frequency Operation

For the unipolar transistor the upper frequency limit is set by the RC time constant, using the junction capacitance and the channel resistance properly averaged over one cycle of operation. The derivation of this value is given in Appendix III. The resulting time constant is

$$\tau = -\frac{l^2}{\mu} \frac{1}{(kV_P - 5)} \log \left[\frac{(1 - \sqrt{k})}{(1 - \sqrt{5/V_P})} \right],$$

where

l = channel length
μ = mobility of majority carriers in the channel region
k = per cent of pinch-off voltage used in the OFF state.

With the $k = 0.94$ corresponding to an OFF resistance of 20 kohms and with a channel length of 0.001″ this becomes

$$\tau = 1.4 \cdot 10^{-9} \text{ sec.}$$

The electric field strength in the channel of an OFF unit reaches values where the carrier mobility starts to become affected. This reduces the high frequency performance[3] by a factor f

$$f = \sqrt{\frac{E_c \cdot l}{V}}$$

where E_c is the critical field where the mobility starts to drop. (For silicon $E_c = 2.5 \cdot 10^3$ volts/cm for electrons, $E_c = 8 \cdot 10^3$ volts/cm for holes). For an n-type channel this will increase τ by less than 50 per cent.

The junction and stray capacitances will introduce additional delay, the time constant of which may be estimated from

$$\tau = RC.$$

The magnitude of the stray capacitances is difficult to estimate. In favorable cases such as shown in Fig. 14(b) where the interconnections are short and direct they should be no larger than the junction capacitances. With $R = 4000$ ohms, $C = 4$ $\mu\mu$F, τ becomes $16 \cdot 10^{-9}$ sec.

Considering also the transients described in the next section the total delay per stage amounts to approximately $3 \cdot 10^{-8}$ sec. This corresponds to an operating frequency of approximately 30 mc/sec. per stage for transient operation. When several units are used in series or parallel the speed decreases correspondingly.

J. Junction Capacitance and Resistance

The influence of the junction resistance and capacitance is to introduce "cross talk." Considering pulse operation this cross talk has to be kept sufficiently low so that it does not interfere with the operation of the logic.

The reverse current of the junction may be as high as 1 μa depending upon surface conditions. At 5 volts bias this would give a junction resistance of 5 megohms. This is unlikely to introduce difficulties as the maximum resistance to ground anywhere in the system is equal to the load resistance, 4000 ohms.

The junction capacitance will allow transients during the rise and fall time of pulses. Their magnitude may be computed as follows. The stored charge in the capacitance $\frac{1}{2}C_{12}$ (see Fig. 5) is

$$Q = CV = \frac{a^2}{2} \sqrt{\frac{\epsilon V}{2\mu\rho}}$$

where ρ is the resistivity in the channel region. With

$a = 0.020''$,
$\rho = 37$ ohm cm and
$\mu = 1900$ cm^2/volt-second
$Q = 4 \cdot 10^{-12} \sqrt{V}$ coulombs.

Differentiation, assuming a rise and fall time of $3 \cdot 10^{-8}$ sec. and ON and OFF voltages of 5 and 15 volts, gives

$$I = \frac{dQ}{dt} = 0.3 \text{ ma.}$$

This corresponds to "spikes" of about 1 volt at the beginning and end of each pulse, having a characteristic time constant of about 10^{-8} sec.

It is interesting to note that this transient reduces the speed of operation through a series chain as shown in Fig. 6 but increases the speed through a chain such as shown in Fig. 7. The fact that an interelectrode capacitance in an electronic device increases the high fre-

quency response is unusual. The conditions for this are of course a phase shift of 180° and a gain of close to unity.

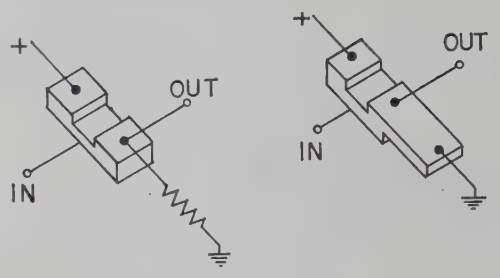
K. Integration of the Resistance

For the simple applications shown here the only passive components used are resistors. Moreover, the resistance values are all similar and generally center between the ON and the OFF resistance of the channel of the unipolar transistor. Therefore, by providing an additional unipolar element its channel could be utilized for the resistance provided that the channel width was properly adjusted. Two alternatives are possible as illustrated by Fig. 8 and Fig. 9. In Fig. 8 the junction has been removed and the resulting resistance is purely ohmic. In Fig. 9 the unipolar element is intact and used as a two-pole with the gate floating. The resulting resistance, then, is nonlinear, as may be seen from Fig. 10, which gives the I_D vs V_D curve of a unipolar transistor with the gate floating. As it happens the nonlinearity is advantageous in that it insures more positive switching action. When the active unipolar transistor in series with the passive one is ON the resistance value of the passive element is high resulting in a more extreme ON voltage for the subsequent stage. When on the other hand the active transistor is OFF the resistance of the passive one is low resulting in a more extreme OFF voltage for the subsequent stage.

The passive unipolar transistor may be combined with the active one as also shown in Fig. 9(b) and may be fabricated simultaneously. The resulting combination may be used as the standard building block of DCUTL. It is superficially different from the block shown in Fig. 4, as the resistor is here on the output rather than on the input side. As in practice, each unipolar transistor has one resistor on its input side and one resistor on its output side; it is immaterial which alternative is chosen for the standard building block. The result is a completely integrated logic system in which all the active and passive components have been built into arrays consisting of semiconductor wafers. An example of a series chain of units similar to Fig. 6 is shown in Fig. 11. Note the complete absence of external circuit elements.

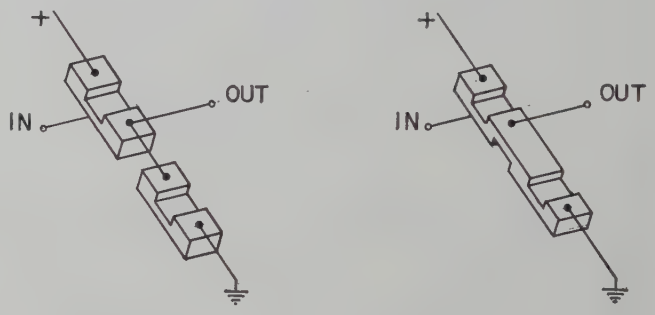
L. Connecting the Elements

By using the technique of making standard elements in arrays, most of the connections between elements are eliminated and replaced by bridges of semiconductor material between the elements joining them. However, a few connections have to be made between different arrays, and also input and output, battery supply, etc., have to be connected. It would be very desirable to have those connections all in one plane so that a simple printed circuit could be used without the necessity of connections back and forth from one side of the wafer to the other. This can be accomplished sometimes by inserting a diode in the few places where such connections are desirable, thus converting the intra-plane con-



(a) (b)

Fig. 8—One stage of DCUTL, (a) transistor plus resistor, (b) transistor with integrated resistance.



(a) (b)

Fig. 9—One stage of DCUTL, (a) transistor and nonlinear resistor (b) transistor with integrated nonlinear resistance.

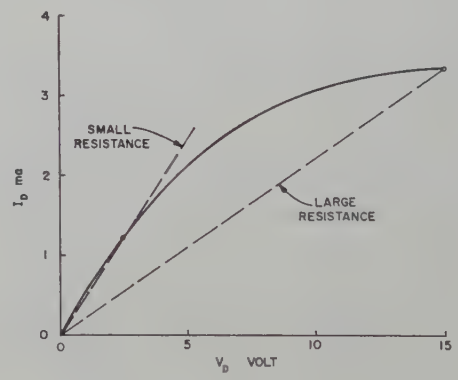


Fig. 10—Drain current vs drain voltage ($V_G=0$).

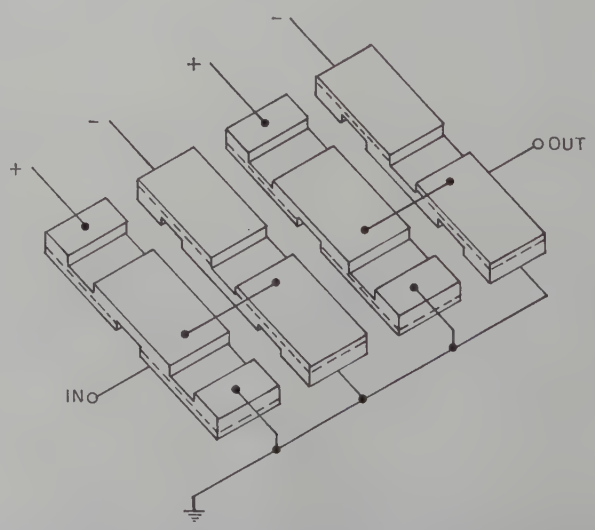


Fig. 11—A series chain of integrated unipolar transistor stages.

nections to in-plane connections. These diodes would always be forward biased. Because of the method of fabrication of the arrays these diodes may be obtained at a negligible extra cost compared to the saving in printed circuit complexity.

III. Integrated Logic Circuits

A. Specific Examples

Some specific examples of integrated logic circuits using DCUTL are shown in the figures below. Each example is shown in logic symbols, in DCUTL, and finally as an integrated device.

Fig. 12 shows a multiple-OR gate, and Fig. 13, a multiple-AND gate. In neither of these are the resistances integrated. The gate load resistances would be supplied by previous stages. The output load resistance could be in the form of an extra unipolar element as shown earlier.

Fig. 14 shows a binary half-adder. This circuit is shown in four versions, the bottom one having the resistances integrated. Other half-adder circuits could be built in a similar manner.

Fig. 15 shows a transfer tree with one input and eight outputs. This tree contains 3×8 unipolar elements operated by three external flip-flops (not shown). In the bottom view are shown the gate connections. Isolation between rows is obtained through the presence of a junction created by a three-step diffusion and lapping operation which has been described by Nelson.[2]

Figs. 16 and 17 show a matrix switch, Fig. 16 with input and output in the same plane, and Fig. 17 with input and output on the opposite side to allow sandwiching to other two-dimensional arrays such as the memory shown in Fig. 18. This memory consists of two unipolar transistors with complementary symmetry connected in a bistable circuit. In this circuit both elements are either ON or OFF simultaneously. Another version is possible in which one element is ON and the other simultaneously OFF. Operation as monostable (memory), bistable or astable multivibrator is possible with modifications of this circuit.

B. Fabrication Techniques

Unipolar transistors have been made in a variety of ways, one of which will be described here as suitable for making large arrays of such units. The main feature of this method is a lapping technique developed by H. Nelson.[2]

Let us consider as an example the multiple-AND gate shown in Fig. 13. The raw material for the device is a wafer of silicon, *n*-type, 37 ohm cm, 0.015-inch thick and about $\frac{1}{2}$-inch square. This wafer is flat-lapped to 0.010 inch. Boron is diffused into the wafer for 20 hours at 1300°C, forming a junction approximately 0.003 inch under the surface. The separating grooves are lapped, taking care to penetrate the junction. The wafer is then soldered onto a printed circuit disk which will be described below. This printed circuit disk provides for

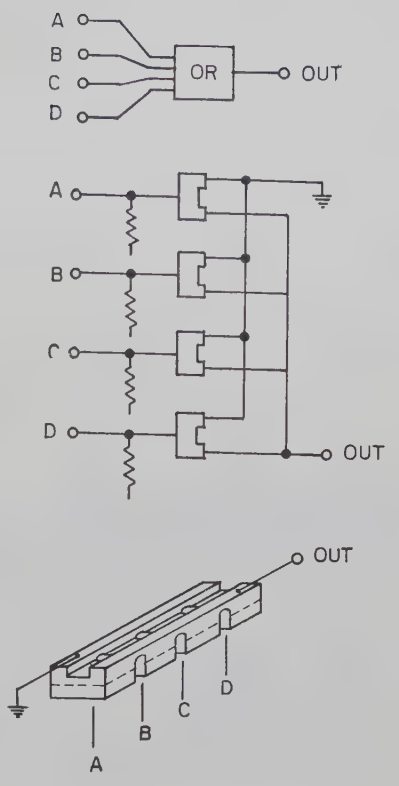


Fig. 12—Multiple-OR gate.

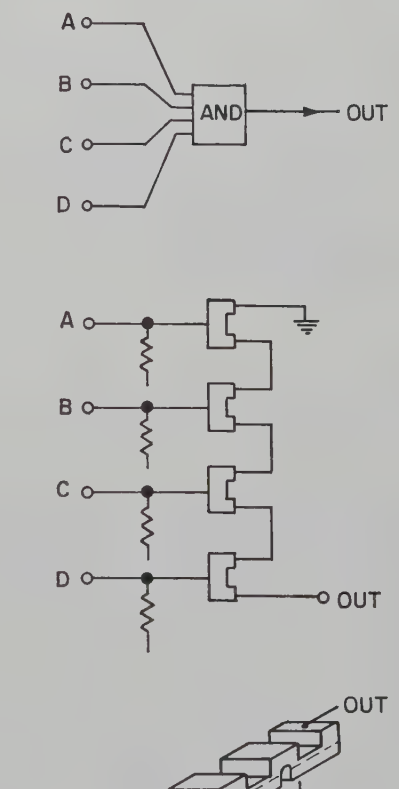


Fig. 13—Multiple-AND gate.

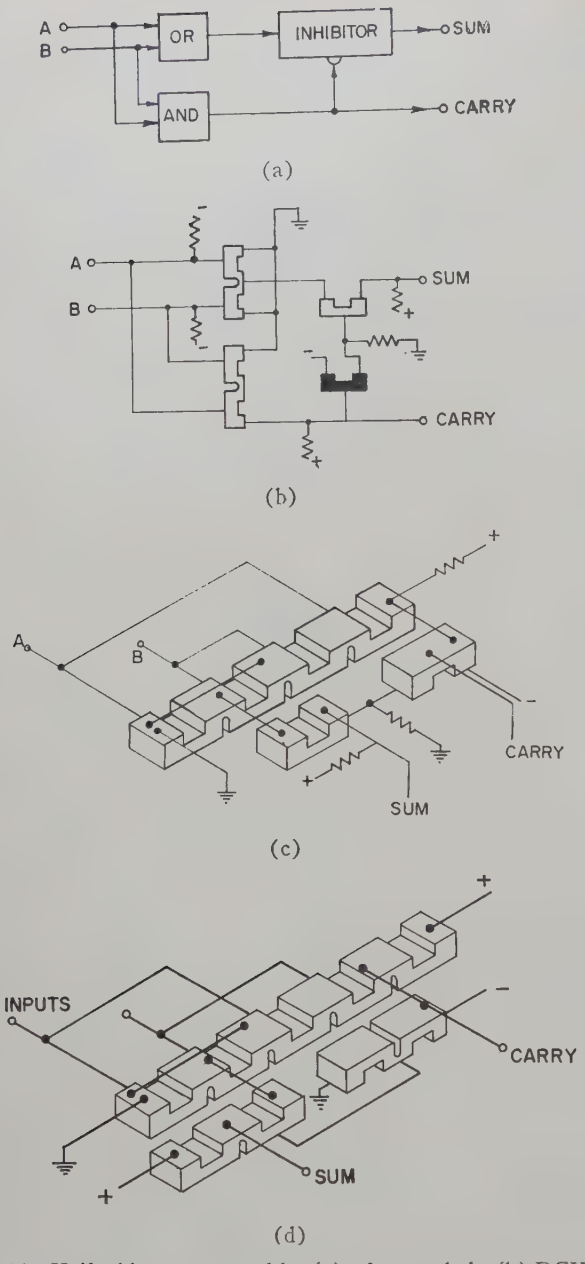


Fig. 14—Half-adder represented by (a) relay symbols, (b) DCUTL, (c) partly integrated DCUTL, (d) integrated DCUTL.

connections to all the gates and at the same time provides the mechanical support necessary to allow the wafer, sandwiched onto the disk, to be flat-lapped down to 0.006 inch. Thereafter the channels are lapped, taking care to leave a channel depth of 0.0005 inch. The wafer is then turned 90°, and grooves are lapped down to the printed circuit metal which is laid bare so that connections later can be made to the printed circuit. Another printed circuit disk is now soldered onto the free surface of the germanium wafer providing connections to the end source and drain. The assembly is finally cut with a diamond wheel into a number of multiple-AND gates. The steps in this procedure are shown schematically in Fig. 19.

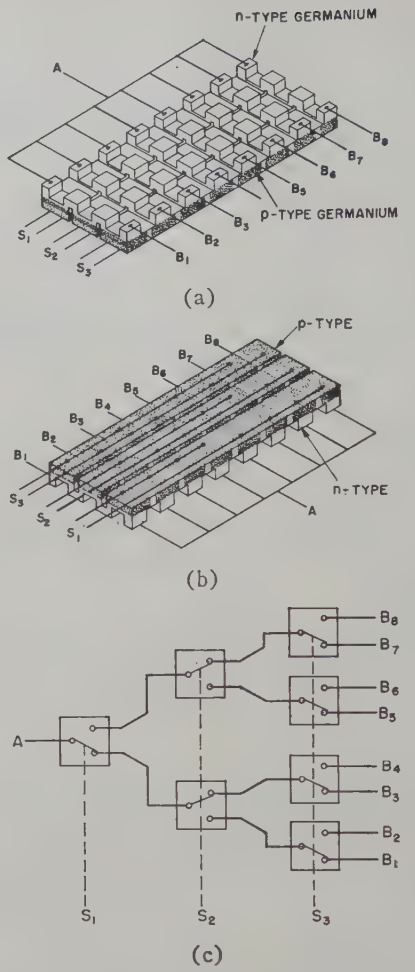


Fig. 15—Three stage tree, (a) top view, (b) bottom view, (c) logical schematic.

C. The Printed Circuit Disk

In order to make contact to all the necessary points in a large array, and at the same time provide mechanical strength, a printed circuit approach appears necessary. This printed circuit may be made on glass of the same thermal expansion coefficient as the semiconductor. The glass disks may be flat-lapped in the same manner as the semiconductor wafers.

A desirable requirement of printed circuit disks is that they allow metallic contacts straight through the disk from one side to the other. If this requirement is fulfilled, great versatility in assembly is obtained in that large circuitry complexes may be assembled in sandwich fashion with layers of semiconductor alternating with layers of printed circuits. As an example, a memory plate such as shown in Fig. 18 may be sandwiched onto a matrix switch plate such as shown in Fig. 17, providing input to the memory. A similar matrix switch plate may be sandwiched onto the other side of the memory to provide output from the memory. Onto the output matrix may be sandwiched another plate, maybe a display panel, an arithmetic unit or another memory plate, etc.

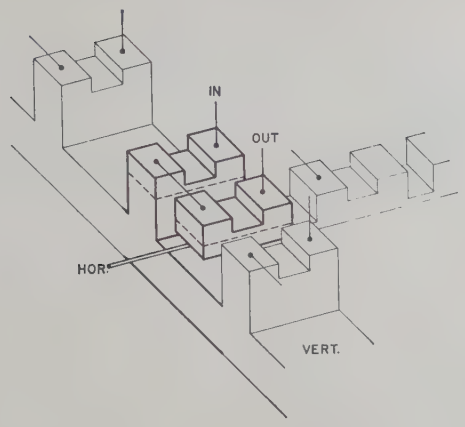


Fig. 16—Matrix switch, input and output on same side.

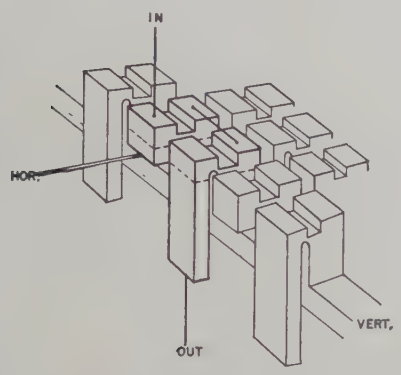
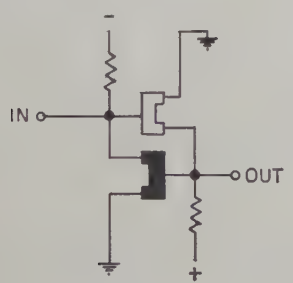


Fig. 17—Matrix switch, input and output on opposite sides.



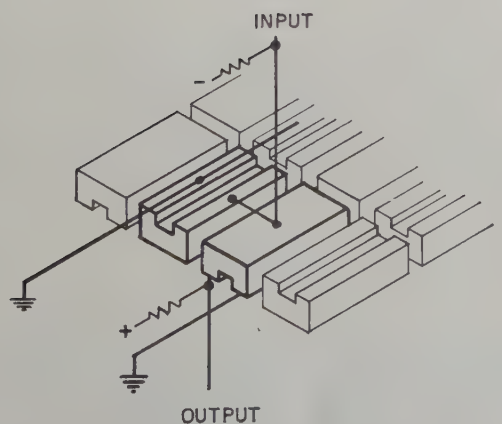


Fig. 18—Memory. One element shown in heavy lines.

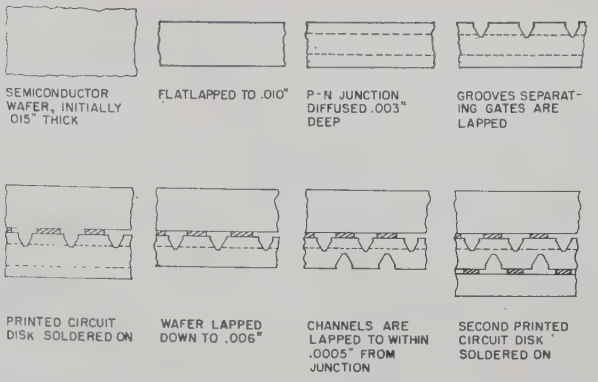


Fig. 19—Manufacturing operations for DCUTL.

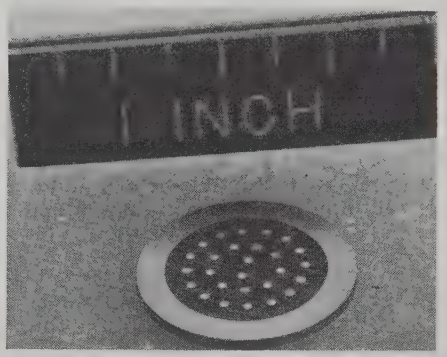


Fig. 20—Glass disk with metallic lead-through conductors.

The manner in which such a printed circuit disk may be made is illustrated in Fig. 20 which shows a glass disk, 0.025-inch thick, with 27 metal leads passing through the glass and arranged in an arbitrary pattern.

The choice of metal-glass construction is dictated by the desirability of hermetic encapsulation of the logic circuitry once it is assembled. Other alternatives, such as using a bakelite printed circuit disk and a separate hermetic encapsulation, are possible, of course.

D. Evaluation of DCUTL

Many of the features of DCUTL such as the direct coupling, the use of a standard building block, etc., are also found in direct-coupled transistor logic (DCTL)[5–7] which is well known. In comparing the two, one finds that the advantages claimed for DCTL—a minimum number of components, no more than two power supplies (one for DCTL) and simple circuitry—also apply to DCUTL. In addition DCUTL offers the following advantages over DCTL.

[5] J. R. Harris, "Direct-coupled transistor logic circuitry," IRE Trans. on Electronic Computers, vol. EC-7, pp. 2–6; March, 1958.

[6] J. W. Easley, "Transistor characteristics for direct-coupled transistor logic circuits," IRE Trans. on Electronic Computers, vol. EC-7, pp. 6–16; March, 1958.

[7] R. H. Beter, W. E. Bradley, R. B. Brown, and M. Rubinoff, "Directly coupled transistor circuits," *Electronics*, vol. 28, p. 132; June, 1955.

1) The signal amplitude is large, approximately 10 volts, and consequently DCUTL can drive, or be driven by, most transistor pulse circuits without additional amplification or voltage division. In contrast DCTL operates with small signals, approximately 0.3 volt. As a consequence DCUTL does not require special care in grounding and is not sensitive to stray signals and noise.

2) Speed of DCUTL is not limited by transistor saturation and is therefore considerably higher than for DCTL. The speed of DCTL can be increased by reducing the transistor base width. Similarly, however, the channel length of the unipolar transistor can be reduced, increasing the speed of DCUTL, so that for similar effort DCUTL is still ahead.

3) There is no trouble with dc levels in DCUTL. In a multiple-AND gate DCTL requires separate bias for each input. In DCUTL the large signals used override any small difference in bias, and the same bias can be used for all inputs.

4) Larger dimensional tolerances, mechanical and electrical, are allowed in DCUTL. The unipolar transistor tolerates high surface recombination velocity, making for stability on life and ease in fabrication. The unipolar transistor has only one junction that has to be good. The only critical dimension of the unipolar transistor, the channel width, can be lapped or etched to correct value by monitoring the channel resistance, a simple procedure.

5) The unipolar transistor offers very good insulation between input and output. This simplifies circuit synthesis in complex networks.

6) Because the drain circuit of the unipolar transistor is purely ohmic (the current passes no junctions), fabrication of two-dimensional arrays of DCUTL is very simple. Nearly all connections may be made in one plane by printed circuit techniques. Very few intra-plane connections are necessary.

Appendix I

Derivation on ON and OFF Levels

Assume that the input gate voltage at an arbitrary stage is V_1. Then according to Dacey and Ross[3] the channel resistance of that stage, assuming an abrupt junction and neglecting the series resistance in the source, is

$$R_{ch} = V_2/I_D \tag{1}$$

where

$$\begin{cases} I_D = \dfrac{1}{l}\left\{J(V_2 - V_G) - J(V_S - V_G)\right\} \\ J(x) = g_0 x\left[1 - \dfrac{2}{3}\sqrt{\dfrac{x}{V_P}}\right] \\ g_0 = \dfrac{2wa}{\rho} \end{cases}$$

a, w and l are shown in Fig. 1, and

V_P = pinch-off voltage for the channel
$V_P = w^2/2\epsilon\mu$.

The gate voltage for the next stage is

$$V_2 = \frac{R_{ch}}{R_{ch} + R} V_B, \tag{2}$$

where

R = gate load resistance

and

V_B = supply voltage.

Setting $V_1 = V_2$, $V_B = V_P$, combining (1) and (2), and inserting practical values, results in

$$V_1 = \begin{cases} -15 \text{ volts (pinch-off)} \\ -5 \text{ volts.} \end{cases}$$

Introducing different values of V_1, and computing V_2 gives the curve in Fig. 21. From this curve it can be seen that any arbitrary deviation from -5 will rapidly converge through a number of stages towards the stable value -5 volts.

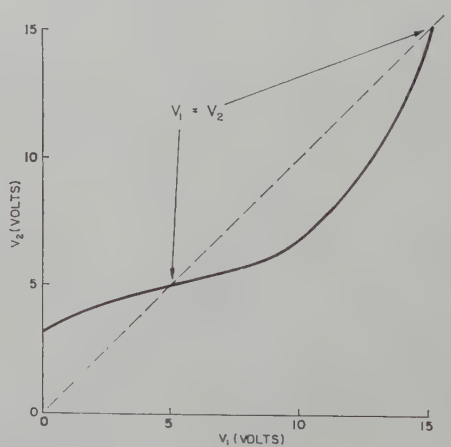


Fig. 21—Output voltage level vs input voltage level.

The solution $V_1 = -15$ represents an unstable condition which will also converge toward -5 volts after a sufficient number of stages. However, by using $V_B > V_P$ and driving past pinch-off even this point may be made stable.

Appendix II

Derivation of AC Stability Condition

Assume a voltage deviation ΔV at the input of an arbitrary stage. Then the output deviation ΔV_2 is

$$\Delta V_2 = \Delta V g_{m\,\mathrm{dyn}} R,$$

where

$g_{m\,\mathrm{dyn}}$ = dynamic transconductance of the unipolar transistor

and

R = drain load resistance.

The dynamic transconductance may be computed as indicated by Dacey and Ross,[3] using the expressions from Appendix I;

$$V_D = V_B - RI_D.$$

Then

$$g_{m\,\mathrm{dyn}} = \frac{dI_D}{dV_G} = \frac{g_m}{1 + Rg_m \dfrac{\sqrt{V_P} - \sqrt{V_D - V_G}}{\sqrt{V_D - V_G} - \sqrt{-V_G}}}$$

where g_m is the static transconductance neglecting feedback in the load resistance

$$g_m = \frac{g_0}{L\sqrt{V_P}} (\sqrt{V_D - V_G} - \sqrt{-V_G}).$$

Operation in the OFF condition, corresponding to pinch-off, gives

$$g_{m\,\mathrm{dyn}} = \frac{g_m}{1 - Rg_m}.$$

The condition that $\Delta V_2 \leqq \Delta V$ gives

$$Rg_{m\,\mathrm{dyn}} \leqq 1.$$

With practical values inserted this condition is easily fulfilled as g_m is very low in the OFF state. In the ON state the corresponding condition is

$$\frac{Rg_m}{1 - Rg_{mp}} \leqq 1$$

where $g_{mp} \approx$ transconductance at pinch-off. This is fulfilled when

$$g_m \leqq 250\ \mu\mathrm{mhos}.$$

Computation of g_m using practical values gives

$$g_m = 100\ \mu\mathrm{mhos}.$$

Appendix III

Derivation of the Time-Constant Expression

The small-signal depletion-layer capacitance of the gate, assuming an abrupt junction and that $V_D = V_S$, is

$$C = \sqrt{\frac{\epsilon}{2\mu\rho V}}.$$

where V is the applied voltage corrected for the junction contact potential. $V = V_{\mathrm{appl}} - \phi$ where ϕ is a fraction of a volt and can usually be neglected. For an n-type channel (50 ohm cm silicon) this becomes

$$C = \frac{2.8\ 10^3}{\sqrt{V}}\ \mu\mu\mathrm{F/cm^2}.$$

For dimensions in Fig. 1 this gives $C = 3.1\ \mu\mu$F at 5 volts.

The resistance of the channel under the same conditions is

$$R_{ch} = \frac{l}{a}\sqrt{\frac{\rho}{2\epsilon\mu}} \frac{1}{\sqrt{V_P} - \sqrt{V}}.$$

For a n-type channel, 50 ohm cm silicon, and with the dimensions of Fig. 1, this becomes

$$R_{ch} = \frac{6.6\ 10^3}{\sqrt{V_P} - \sqrt{V}}.$$

The RC product then becomes

$$RC = \frac{l^2}{2\mu} \cdot \frac{1}{\sqrt{V_P V} - V}.$$

Averaged over one cycle, assuming that the largest negative gate voltage is kV_P, we have the RC time constant

$$\tau = \frac{l^2}{(kV_P - 5)2\mu} \int_5^{kV_P} \frac{dV}{\sqrt{V_P V} - V}$$

$$= -\frac{l^2}{(kV_P - 5)\mu} \log \frac{(1 - \sqrt{k})}{1 - \sqrt{\dfrac{5}{V_P}}}.$$

This relation is shown in Fig. 22. It is similar to expressions derived by Shockley,[4] and others,[3] but extended towards the "expop" region. For very large k the expression is not valid.

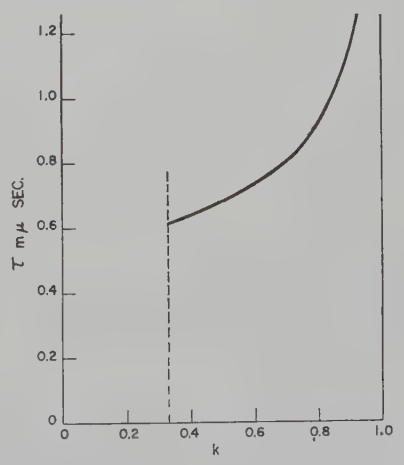


Fig. 22—Time constant τ vs relative off-voltage k.

Acknowledgment

The writers are indebted to H. Nelson for pioneering work on this type of unipolar transistor and methods of making them, and to J. Briggs and R. R. Vannozzi for help with fabrication and measurements.

P-N-Π-N Triode Switching Applications*

V. H. GRINICH† AND I. HAAS†

Summary—The characteristics and switching applications of a developmental diffused silicon *p-n-π-n* triode are discussed. Although this unit is at present in a two-watt package, it is capable of handling short pulses of current of the order of 100 amperes. The electrical characteristics which consist of a low and high conductivity region (over 500 megohms and less than 1 ohm respectively), with an intermediate negative resistance region, are controllable by the base lead, and hence make it a flexible device for applications in the computer and communications fields. The theoretical and practical limitations are discussed. Experimental data covering current handling capabilities, frequency limitations and switching times are presented in conjunction with representative circuits. Two particular circuits discussed are an 80-ampere 500-mμsec pulse generator with rise and fall times in the order of 150 mμsec and 300 mμsec respectively, that can operate up to a kilocycle repetition rate, and a 4-ampere 60-mμsec pulse generator with a PRF of 100 kc. Other examples described include monostable, bistable, and astable circuits, as well as types of communication circuitry for a wide range of currents.

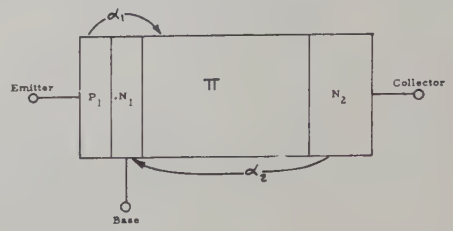
Introduction

THE *p-n-π-n* triode is a silicon three-terminal switching device incorporating the basic principles of junction transistors and approaching the requirements of the ideal switch. By making a thick near-intrinsic π[1] base region[2] one obtains a greater stand-off voltage[3] than is possible in the more conventional *p-n-p-n* triode. As in *p-i-n* diodes[4] this intrinsic layer does not add significantly to the ON resistance since it is heavily conductivity modulated. For high-speed turn-off, design compromises are necessary because of the additional charge stored in the π region. The device to be discussed here is enclosed in a two-watt package.

Physical Characteristics

The *p-n-π-n* dc characteristics can be analyzed on the basis that it is composed of two junction transistors,[5–10] the collector of one being common to the base of the other and vice-versa.

When a reverse bias (negative with respect to the emitter) is applied to the collector n_2 (we define the collector contact as the ohmic contact to the region n_2) all of the junctions except $n_1\pi$ are forward biased (see Fig. 1). As the current through the device is increased,



$$E = \rho J$$

$$\alpha_2 = e^{J/J_0}\left[\cosh\sqrt{(J/J_0)^2 + (W/L_n)^2} + \frac{\sinh\sqrt{(J/J_0)^2 + (W/L_n)^2}}{\sqrt{1 + (W/L_n)^2(J_0/J)^2}}\right]^{-1}$$

where $J_0 = 2kT/q\rho W$.

For

$$J/J_0 \ll 1, \qquad \alpha_2 = \operatorname{sech} W/L_n, \qquad \text{(diffusion)}.$$

For

$$J/J_0 \gg 1, \qquad \alpha_2 = e^{-\tau_0/\tau} \qquad \text{(drift)}$$

where $\tau_0 = W/\mu_n\rho J$.

Fig. 1—Current transfer-ratio of *n-π-n* section.

the *V-I* relationship is primarily determined by the $n_1\pi$ junction reverse characteristics. The current transfer ratio of the *n-π-n*, α_2, remains fairly small at first, but then rises gradually toward unity. When the current density in the π region rises to a value such that the sum of the alphas of the *p-n-π* and *n-π-n* reach unity, the unit becomes regenerative,[5,6] and the current through the device is only limited by the external circuitry.[7] For increasing base currents, the *V-I* characteristics are identical with those of a transistor for base current control, as long as the current density J in the π region is lower than the value required to increase the sum of the alphas to unity, as shown in Fig. 2. When the unit has switched to the ON state, electrons are injected into the π region. This raises the effective conductivity of the π region and, hence, gives a very low saturation resistance. The effective saturation resistance of these devices is in the order of one ohm.

* Manuscript received by the PGEC, February 12, 1959; revised manuscript received, April 13, 1959. This paper was presented at the 1959 Solid-State Circuits Conference, Philadelphia, Pa., February 12–13, 1959.

† Fairchild Semiconductor Corp., Palo Alto, Calif.

[1] M. B. Prince, "Diffused *p-n* junction silicon rectifiers," *Bell Sys. Tech. J.*, vol. 35, pp. 661–685; May, 1956.

[2] The "π" refers to the thick base's bulk resistivity type and was chosen here to emphasize that the bulk is a lightly doped *p*-type. See *ibid.*

[3] The stand-off voltage in this device is the maximum voltage that can be applied between collector and emitter without causing it to go into the negative resistance region.

[4] D. A. Kleinman, "The forward characteristic of the *p-i-n* diode," *Bell Sys. Tech. J.*, vol. 35, pp. 685–707; May, 1956.

[5] J. L. Moll, M. Tanenbaum, J. M. Goldey, and N. Holonyak, "*P-n-p-n* transistor switches," PROC. IRE, vol. 44, pp. 1174–1182; September, 1956.

[6] J. A. Hoerni and R. N. Noyce, "*P-n-π-n* switches," 1958 IRE WESCON CONVENTION RECORD, pt. 3, pp. 172–175.

[7] R. W. Aldrich and N. Holonyak, "Multiterminal *p-n-p-n* switches," PROC. IRE, vol. 46, pp. 1236–1239; June, 1958.

[8] J. T. Nelson, J. E. Iwersen and F. Keywell, "A five-watt ten-megacycle transistor," PROC. IRE, vol. 46, pp. 1209–1215; June, 1958.

[9] C. W. Mueller and J. Hilibrand, "The thyristor—a new high-speed switching transistor." IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 2–5; January, 1958.

[10] I. M. Mackintosh, "Three-terminal *P-N-P-N* transistor switches," IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 10–12; January, 1958.

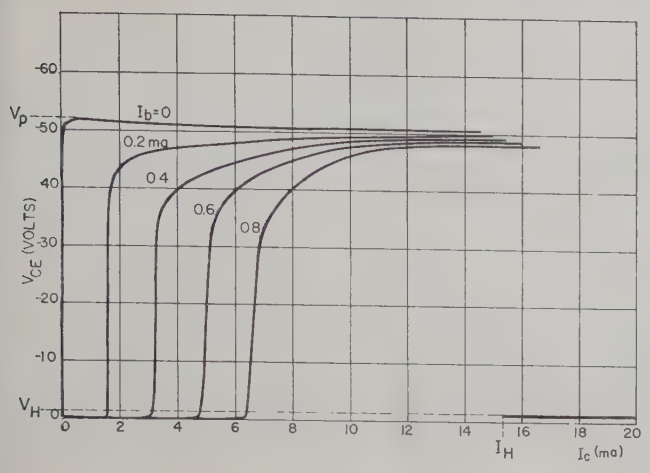


Fig. 2—*P-n-π-n V-I* plot.

A unit made of 100 ohm-cm material may have a saturation resistance of an ohm, while that of a 10 ohm-cm unit will be 0.5 ohm. The impedance in the low conductivity region is about 500 megohms.

The magnitude of the switching current, *i.e.*, the current at which the sum of the alphas is unity, varies inversely as the π layer's resistivity, and varies as the ratio of the base width to the diffusion length, assuming unity injection efficiency at the π-n_2 junction. The switching current will increase as the injection efficiency decreases, since α_2 decreases. It is thus possible to design for a given switching current.[6] Avalanche triodes and *p-n-π-n*'s differ in that the high collector spreading resistance, $r_{cc'}$,[8] of the latter during switching leaves M, the multiplication factor, practically equal to unity. The high field across the π region makes the transit time in this layer very small. The transit time is limited by terminal velocity effects.

Circuit Applications

The bistability of the device makes it applicable to bistable, monostable or astable applications. The pulse current handling capacity of the device is practically unlimited, as compared to present day transistors, because of the high conductivity in the ON state.

The parameters to be used in this discussion are the following: the holding voltage, V_H, which is the collector to emitter voltage of the device in the ON state at the current I_H; the holding current, I_H, which is the current at which the device goes from the ON to the OFF state; R_S, which is the incremental resistance in the ON state; the breakdown voltage, V_p, the peak voltage this device can stand off in the OFF or low conductivity state. The effects of these parameters on circuit performance will be discussed in the subsequent paragraphs.

Bistable Circuits

The operation of bistable circuits using *p-n-π-n*'s is similar to that of regular transistors. The base current required to turn on the device is primarily determined by the α of the *p-n-π* and the holding current I_H, since the latter is determined by α_2. Once the device has been turned on, of course, it will act as a memory or flip-flop and does not require an input to keep it on.

The basic circuit is shown in Fig. 3. The rise time is affected by the input in the same manner as transistors. The curves shown here were taken with minimum drive at the base. The effect of using higher collector supply voltages in decreasing the transit time is clearly seen in the rise-time variation. In this case, the switching was entirely controlled from the base.

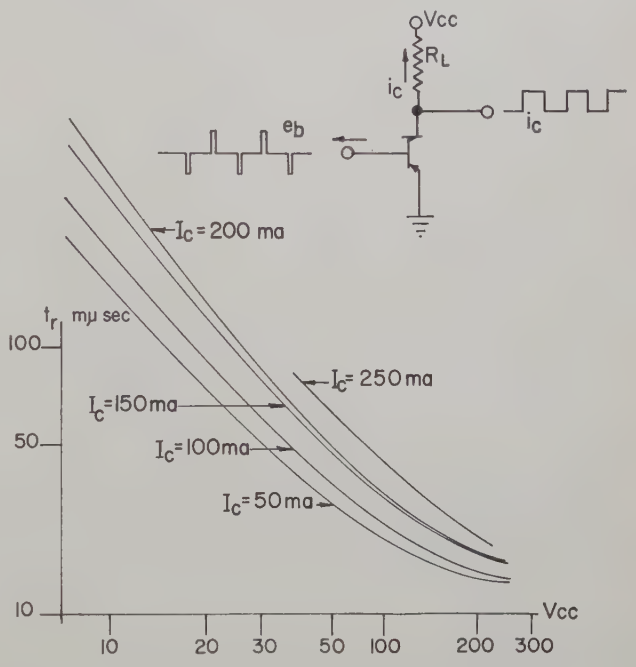


Fig. 3—High-current storage circuit.

For the particular device under investigation, 250–300 ma was the maximum switchable current that could be controlled from the base. The upper threshold of base-controlled current may be due to several reasons. One theory suggested in the literature[6] for the turn-off mechanism is that the base turn-off current is about the same as the collector current.

If such is the case, then the maximum switchable collector current might depend on the base-spreading resistance. If the product of the collector current and the base-spreading resistance is greater than the base to emitter breakdown voltage, then the emitter junction breakdown limits the region of control from the base. Experimental results show this general tendency, but the distributed nature of the current path makes it difficult to get an accurate and simple model to verify this exactly.

Fig. 4 shows the variation of fall and storage times with collector current. The fall time lies between 3 and 4 μsec, and the storage time is about 1.5 μsec. Typical units with lower peak voltages had the following switching times at 200 ma collector current: $t_r = 50$ mμsec, $t_s + t_f = 120$ mμsec, at $I_{b1} = I_{b2} = 50$ ma. t_d was about

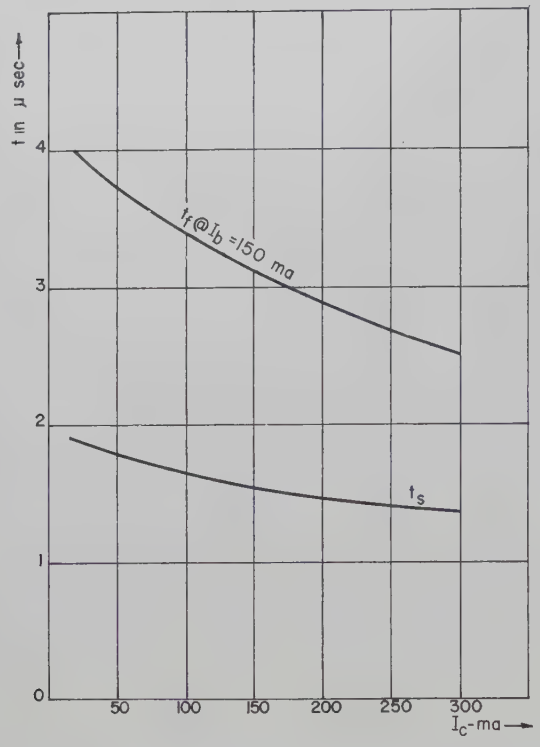


Fig. 4—Storage and f all times as a function of collector current.

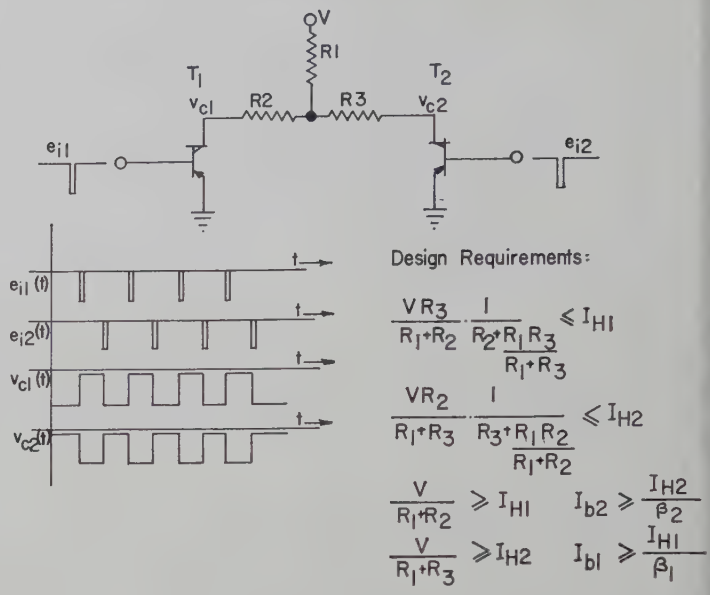


Fig. 5—Two-phase pulse former or flip-flop.

15 mμsec, measured between 10 per cent points of input and output signals. A few of these experimental units were made to switch about 900 ma while controlled entirely from the base. The switching times of this unit were of the same order of magnitude as the previously discussed cases.

One way to get around the difficulty of loss of base control is by employing current steering techniques. Fig. 5 shows this basic circuit and requirements. The resistances are chosen so that if one of the units is off and the other turned on, the available collector current is greater than the holding current I_H; if the first transistor is now turned on, the current available to the second should be smaller than its holding current. Thus, if negative voltage or current pulses are alternately applied to the two bases, the output at one of the collectors will be a pulse whose width is equal to the duration between the leading edges of the trigger pulses, while the output at the other collector will be the complement of the first, as shown in the figure. As can be seen from the equations, the maximum switchable current is proportional to the holding current I_H. The advantage of this circuit over the single stage is mainly that the base drive is of one polarity only, and the base drive magnitudes are determined using the same considerations as in conventional transistors. Furthermore, it also provides all the advantages of a flip-flop. Rise and fall times of this circuit are in the order of 0.2 μsec. The units used in this circuit had peak voltages of about 60 volts. Diode logic can be performed at one of the inputs. The other input is used as the resetting element, and either of the collectors can be used as the read-out.

Monostable Circuits

The basic monostable circuit is shown in Fig. 6. The static load-line, as determined by V and R_1, is chosen so that $V/R_1 < I_H$. When a negative current pulse of magnitude greater than I_H/β is applied to the base, the device is turned on. The dynamic load-line is now determined by R_2, C and the dynamic impedance of the device. The capacitor thus discharges, and a current, whose magnitude is determined by R_2, the dynamic saturation impedance and the switching time, flows through the load R_2. Fig. 7 shows the instantaneous voltage and current of such a circuit.

In the case of a perfect switch, the output would consist of a very fast jump to a high current, the latter decreasing at a rate determined by the circuitry's time-constant. However, in the p-n-π-n the space charge in the π region spreads gradually, thus going through a state of premature saturation before reaching full conductivity modulation. The current rise and magnitude is thus limited by this effect, which also exhibits itself in the collector voltage rise.

Fig. 8 shows the high speed V-I characteristics of a low breakdown device. This curve was taken for a circuit as shown in Fig. 6. The time marks indicate the instantaneous operating points, using the leading edge of the turn-on pulse as the reference. The effect of premature saturation is observed here in that it takes a finite length of time for full conductivity modulation of the π region to set in. During this interval the collector to emitter voltage is higher than in the dc case as seen in Fig. 2. It can be seen that the collector becomes forward biased during the last part of the cycle. The

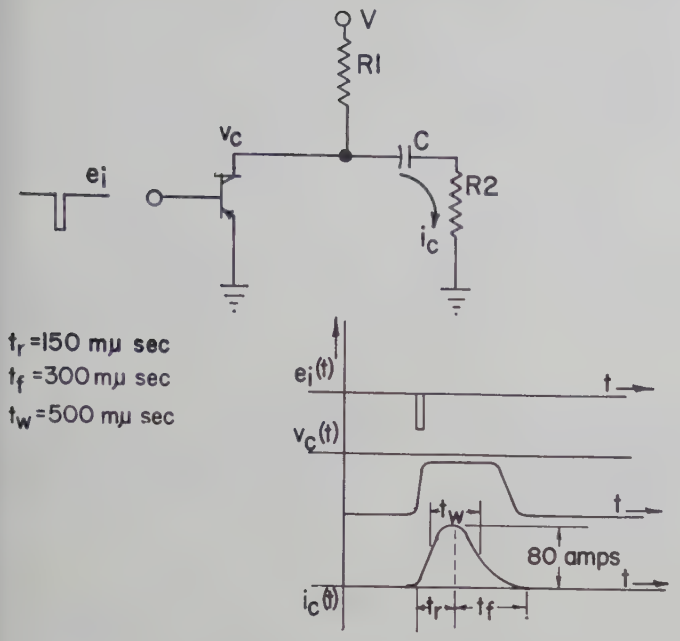


Fig. 6—High-current monostable circuit.

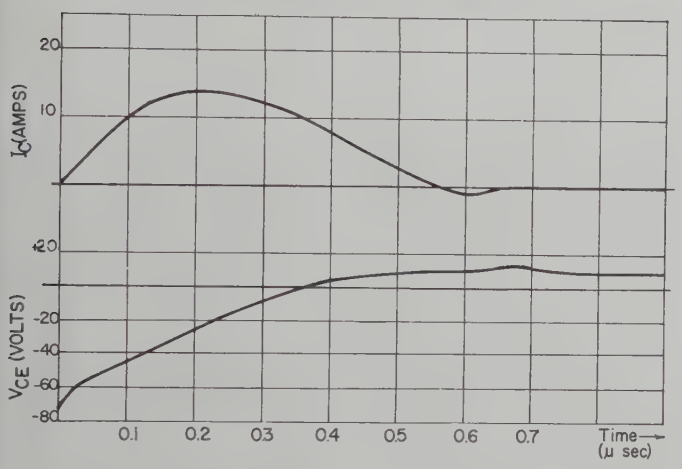


Fig. 7—Instantaneous collector current and voltage waveforms of high-current monostable circuit shown in Fig. 6.

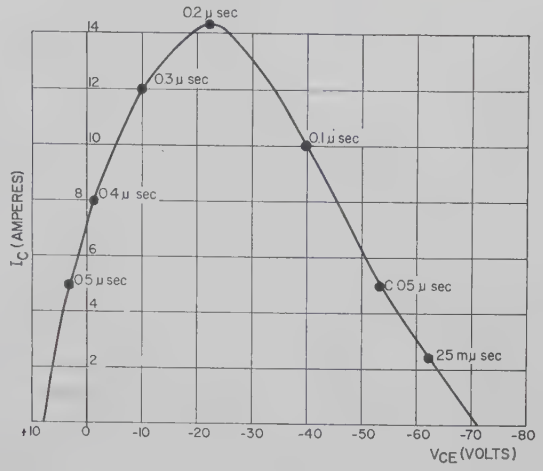


Fig. 8—High-speed *V-I* plot. (Sweep Time: 0.8 μsec.)

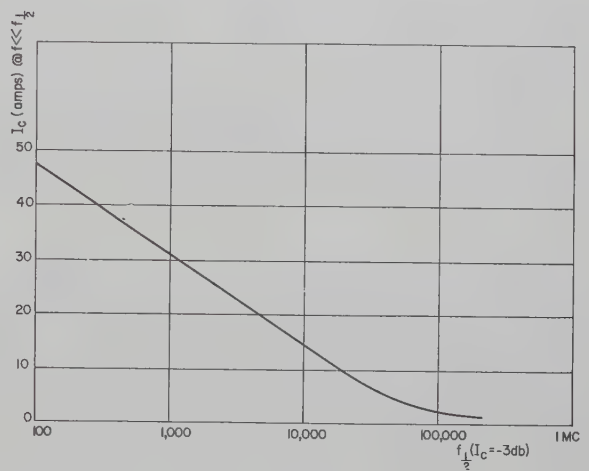


Fig. 9—Current handling capacity of *p-n-π-n* as a function of frequency in the circuit shown in Fig. 6.

reason for this is not known. The time marks on this figure identify the locus of the intersection of the dynamic load-line and the *V-I* characteristics. Using high voltage units, *i.e.*, with a breakdown voltage, V_p, of 200–300 volts, it is possible to obtain very large current pulses. Such high voltage units have been used to switch currents as large as 100 amperes.

To a first approximation, the peak current can be estimated from the breakdown voltage, V_p, and the collector saturation resistance, which in this case is about one ohm. As a rule of thumb, the R_2C time constant, therefore, should be at least as long as the expected rise-time. Thus letting R_2 equal one ohm, and V_{CC} 100 volts, the peak current will be about 50 amperes, if C is chosen to be at least 0.2 μF. Using these high voltage units, 80-ampere current pulses with a rise-time of about 150 mμsec at a pulse repetition frequency of one kc have been obtained with the circuit shown in Fig. 6. The power handling capability of this device is determined by the rms power dissipation.

Fig. 9 shows the current-frequency relationship of a low voltage unit (about 50 volts), with no heat sink. The current was measured at a frequency below that at which the peak current becomes frequency sensitive. The peak current depends strongly on the amount of charge pumped into the large base, since the rise-time depends on the rate at which the minority charge carriers can build up in the wide π region.

Using circuits similar to that shown in Fig. 6, the maximum pulse repetition frequency at low frequencies is first limited by the desired rms dissipation. Aside from this, it is mainly determined by the R_1C time constant at low frequencies. For optimization purposes, this time can be reduced considerably by clamping the collector to a voltage smaller than V_{CC} (see Fig. 6) through a diode. This could also be simulated by biasing the device in the asymptotic region of breakdown, as long as the rated dissipation is not exceeded. At higher fre-

quencies, when the magnitude of R_1C becomes comparable with the fall or recovery time of the device, the latter becomes the limiting factor (see Fig. 4). Thus, going to the limit, it has been possible to obtain 4-ampere pulses at 100 kc, with rise and fall times of 20 and 40 mμsec, respectively, and a pulse of 60 mμsec. An air-blast was used in this case for cooling.

An almost identical circuit, shown in Fig. 10, can be used for inductive switching. The basic operation of this circuit is similar to that previously discussed. Rise-time and peak current, however, now have to be considered in unison. R_2 must be large to approach a current source, but also decreases the peak current proportionally. For a 16 μh inductive load and an R_2 of 100 ohms, an output current pulse of about 400 ma at a rise- and fall-time of 35 to 75 mμsec, respectively, can be achieved. Trading speed for peak current, however, one can obtain the same current amplitudes as in the case of the noninductive load. The collector rise-time is in the order of 20 mμsec because of shunt peaking. One additional requirement on the base pulse now is that its width should be greater than the time required for the current in the load to rise above the holding current, I_H.

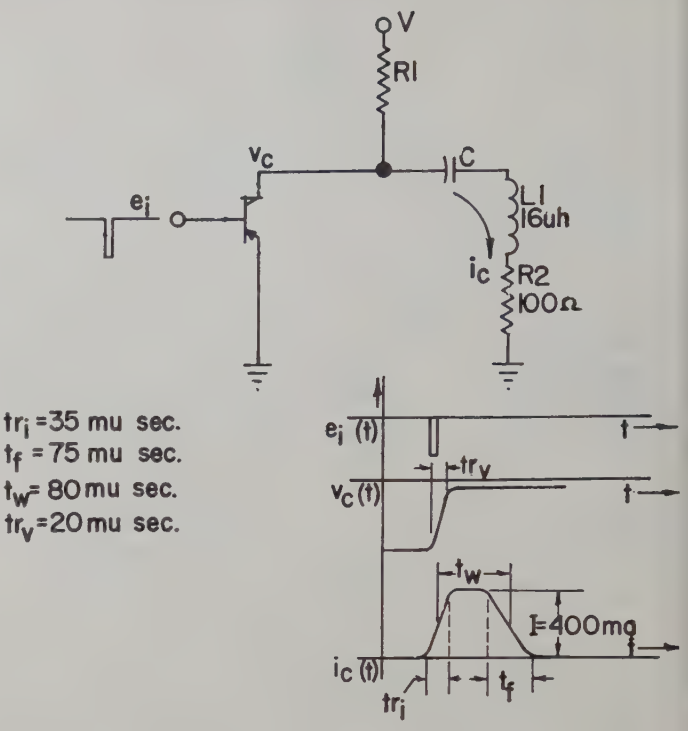


Fig. 10—Monstable memory core switch.

Astable Circuits

The simplest astable circuit would be that shown in Fig. 6 with the base floating and V raised such that the load-line due to R_1 intersects the low frequency V-I characteristics in the negative resistance region. The frequency, current and power considerations, are the same as discussed for the above-mentioned circuits. This mode of operation can be used for pulse and frequency modulation applications at high peak currents by controlling the base. Thus, if a signal whose amplitude is not large enough to drive the device into saturation is applied to the input, the output, either collector current or voltage, will be frequency modulated if the device is never driven into saturation. On the other hand, the latter effect may be used for pulse modulation, as the device will be driven out of oscillation when saturated.

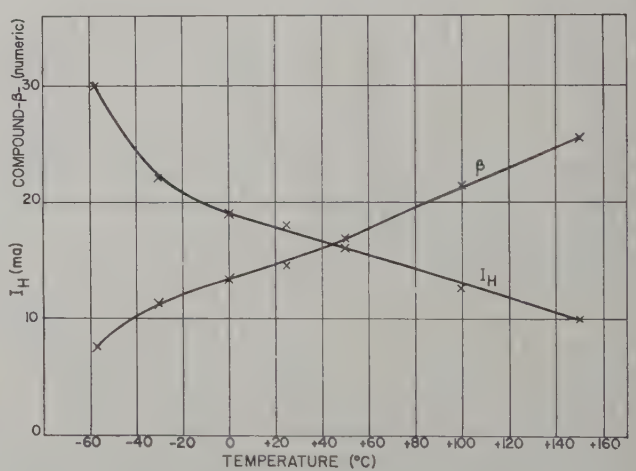


Fig. 11—Temperature dependence of compound β and holding current I_H.

General Discussion

The p-n-π-n is practically noise insensitive when operated as a three-terminal device, since the operating point need not be chosen close to the breakdown voltage as in the two-terminal case for voltage control. The jitter was monitored on an Edgerton TW oscilloscope whose resolution is better than 0.1 mμsec, but was found to be too small to be measured. Rise times are practically unaffected by temperature between -55°C and $+150$°C.

The three-terminal device is also less sensitive to firing by steep voltage wave fronts applied across collector to emitter, since the normally low source impedance in the base will return the capacitive current to ground. Similarly, the effects of high temperature I_{CO}, will be greatly minimized if a reasonably low dc resistance exists between base and emitter. The temperature dependence of the holding current and compound beta was measured and is shown in Fig. 11. The worst case design considerations are identical with that of transistors, except for one additional parameter, namely, the holding current, I_H.

The choice of device parameters depends mainly on the application intended. A high breakdown voltage, V_p, is desirable in all cases considered here, except in astable circuits when the frequency varies inversely as

V_p. For circuits as shown in Figs. 6 and 9, a high holding current I_H is desirable for high frequency applications. In current steering techniques of the type presented here, the magnitude of the holding current is dictated entirely by the circuitry.

Conclusion

The advantageous features gained by using the p-n-π-n as compared to similar two- and three-terminal devices[9–12] can be summarized in Table I.

TABLE I

	Advantages	Disadvantages
Triode over Diode	1) base control 2) not susceptible to dv_c/dt and I_{CO} firing	One additional terminal
p-n-π-n over p-n-p-n	Independent design control over voltage and holding current with no sacrifice in R_s	Some compromise in speed

High voltage p-n-π-n triodes can be used as an excellent approximation of a perfect switch for a large number of applications. The speeds and current requirements in coincident current core-switching are readily met by existing experimental devices. A variety of other high current applications can be envisioned.

[11] J. Philips and H. C. Chang, "Germanium power switching devices," IRE Trans. on Electron Devices, vol. ED-5, pp. 13–18; January, 1958.
[12] D. K. Bisson, "A medium power silicon controlled rectifier," 1958 IRE WESCON Convention Record, pt. 3, pp. 166–171.

An Electro-Optical Shift Register*

T. E. BRAY†

Summary—**An electro-optical shift register composed only of electroluminescent (EL) and photoconductive (PC) cells was designed and successfully operated. While its measured operating speed probably does not make this shift register currently competitive in high-speed applications, it is amenable to construction in an extremely small volume, and has certain other unique characteristics.**

Introduction

TO facilitate an understanding of the circuitry of an electro-optical shift register, a brief discussion of electroluminescence and photoconduction is in order. Materials which emit light when an electric potential is properly applied, are termed *electroluminescent*. Examples are ZnS: (Cu, Ag). While phosphors responsive to dc have been found, only ac responsive phosphors are presently efficient enough to be useful. *Photoconductors* (of the type considered here) are essentially light-sensitive resistors. Photons absorbed by photoconductors excite charge carriers, which results in an apparent decrease of resistance with increasing illumination.

It has been known for some time that EL and PC materials may be combined electrically and optically to perform amplification and switching functions. The so-called light amplifiers and storage panels often utilize EL and PC materials. Electrical amplifiers and switching circuits of various types have been proposed and reported.[1–4]

An electro-optical shift register was designed, since a shift register provides a relatively good basis for the evaluation of new logic and memory techniques. Other electro-optical shift registers have been proposed.[1,2] This register differs from those with which the author is familiar.

The Basic Regenerative Element

The basic element of this shift register is the regenerative connection shown in Fig. 1(a). A thorough under-

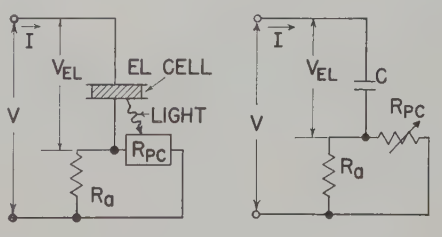


Fig. 1—Basic regenerative connection (a) and its electrical-equivalent circuit (b).

* Manuscript received by the PGEC, March 6, 1959; revised manuscript received April 10, 1959. This paper was presented at the 1959 Solid-State Circuits Conference, Philadelphia, Pa., February 12–13, 1959.
† Electronics Lab., General Electric Co., Syracuse, N. Y.

[1] E. E. Loebner, "Opto-electronic devices and networks," Proc. IRE, vol. 43, pp. 1897–1906; December, 1955.
[2] S. K. Ghandi, "Photoelectronic circuits," Proc. IRE, vol. 47, pp. 1–11; January, 1959.
[3] C. F. Spitzer, "Lumistors: Applications and Limitations," *Proc. NEC*, vol. 14, pp. 353–362; 1958. Or, *Electronic Equip. Engrg.*, vol. 7, pp. 34–38; February, 1959.
[4] B. O. Marshall, Jr., J. R. Bowman, and F. A. Schwertz, "Computer Components," Mellon Inst. Industrial Research, Fellowship No. 347, Quart. Rept. Nos. 5 and 6; 1952.

standing of this EL-PC combination is necessary to an understanding of the shift register. An optical feedback path from EL to PC allows the possibility of bistability. Fig. 1(b) shows an electrical equivalent circuit. The EL cell is predominantly capacitive; resistor R_a is present to allow changes in the driving point characteristics to be made. (In the actual shift register, resistor R_a is replaced by a PC illuminated by the previous-stage EL cell.) For small ac input voltages, the input current will be small because of the high PC resistance. As the input voltage is further increased, sufficient light is emitted to noticeably lower the PC resistance, increasing the current and light output. Due to the positive feedback, the circuit continues this regenerative action until equilibrium is reached. Further increases in applied voltage merely serve to increase the light output and circuit current nearly in proportion to the applied voltage.

Consider the graph of the rms driving point voltage and current shown in Fig. 2. The bistable mode is possible with the familiar S-shaped characteristic. If R_a becomes too small, bistability is no longer possible. Hence varying R_a allows control of the driving point characteristic.

Analysis

An analysis of this regenerative circuit was made to aid in understanding the operation of the shift register. The analytical results obtained were based on the following assumptions.

1) The EL cell is purely capacitive. This assumption is well satisfied in practice.
2) The excitation voltage is sinusoidal. This was true for the work reported here.
3) PC responds only to the time-averaged brightness from the EL cell. It is not obvious that this is true; however, the excitation (carrier) period was chosen sufficiently shorter than the PC response time to allow this assumption to be met.
4) PC is a linear resistance; the volt-ampere relation is a straight line through the origin at any frequency. Fig. 3 shows the PC characteristics, with close agreement between actual and assumed shapes.
5) The transfer characteristics may be written

$$R_{PC} = aV_{EL}^{-\eta}, \qquad (1)$$

where a and η are empirically determined constants.

This assumed relation between the apparent photoconductor resistance, R_{PC}, and the rms voltage applied to an EL cell, V_{EL}, in optical proximity, is of paramount importance. Fig. 4 shows the actual and assumed transfer characteristic indicating a relatively good comparison.

The analysis is given in the Appendix. Note that the results involve the steady state V-I characteristics, and do not account for PC time-constant effects.

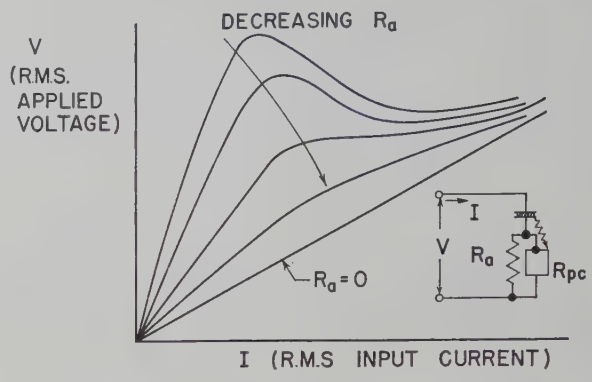


Fig. 2—Typical driving-point characteristics of regenerative electro-optical circuit.

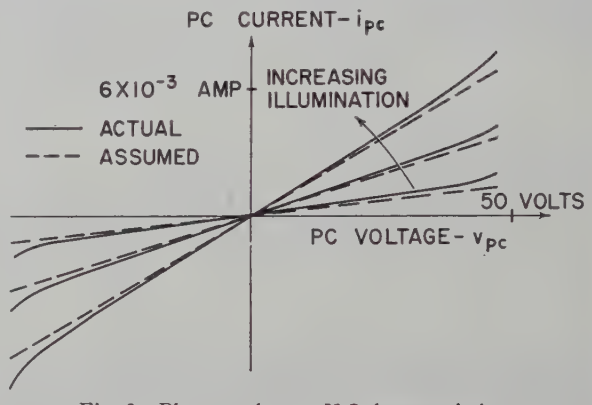


Fig. 3—Photoconductor V-I characteristics.

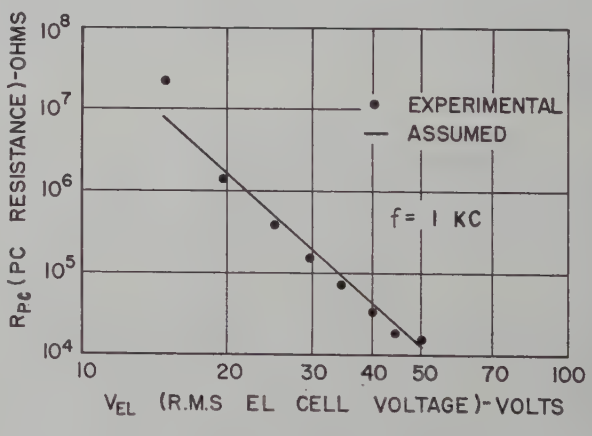


Fig. 4—EL-PC transfer characteristic.

Following are some major results of the analysis.

1) The relation between the rms voltage applied to the terminals of the regenerative connection and the rms current is

$$V = IX_{EL}[1 + (aX_{EL}^{\eta+1}I^{\eta} + \beta)^{-2}]^{1/2} \qquad (2)$$

In this equation,

X_{EL} is the reactance of the EL cell at the excitation frequency,
V and I are, respectively, the terminal rms voltage and current,
a and η are constants [see (1)],
β is a parameter given by the ratio of X_{EL} to the parallel resistance R_a.

The V-I characteristic of the regenerative connection is strongly influenced by the value of β.

2) The limiting value of a critical ratio $X_{EL}/R_a=\beta$ for bistability may be determined as a function of the exponent η from (3). This relation is shown graphically in Fig. 5.

$$\beta_0 = 2/\sqrt{27}\,\frac{(\eta-1)^{3/2}}{\eta}. \tag{3}$$

Values of β in the forbidden region yield V-I curves which are not multivalued, and hence are not bistable.

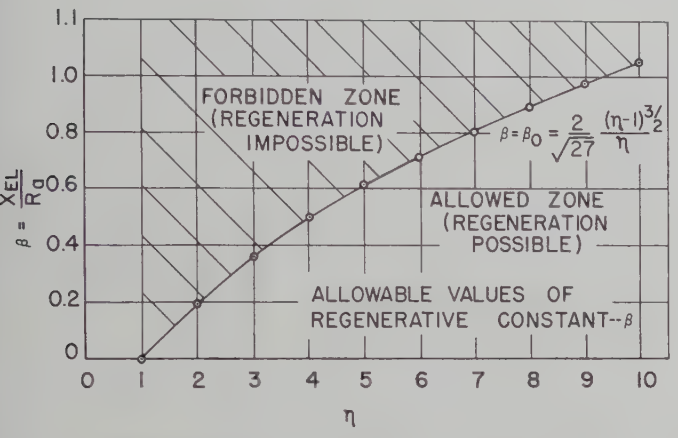


Fig. 5—Allowable values of regenerative constant $\beta=X_{EL}/R_a$.

Other relations which give the peak and valley voltages and currents are given in the Appendix. Experimental confirmation of the analytical results is shown in Fig. 6. While the discrepancy is not negligible, the results are useful as design criteria.

Final Circuit

It was deemed desirable that the shift register not depend on accurately controlled time constants in the PC's; the resulting configuration is shown in Fig. 7. Three regenerative connections (of the type previously considered) are required per bit of information. These are referred to as elements A, B, and C in the figure. Excitation voltages are applied in time sequence as shown. The information state may be considered to be the optical condition of those regenerative connections which are being excited electrically. Three elements per bit are required to prevent spurious information propagation; the connection not excited provides optical buffering.

Shifting takes place as follows. Assume elements A contain the information (are excited electrically). The optical output from each A element influences the driving-point characteristic of each B element. When the B elements are excited, the condition each B element assumes depends on the previous state of the adjacent A element.

This can be more easily understood by considering only two elements in the register. (See Fig. 8.)

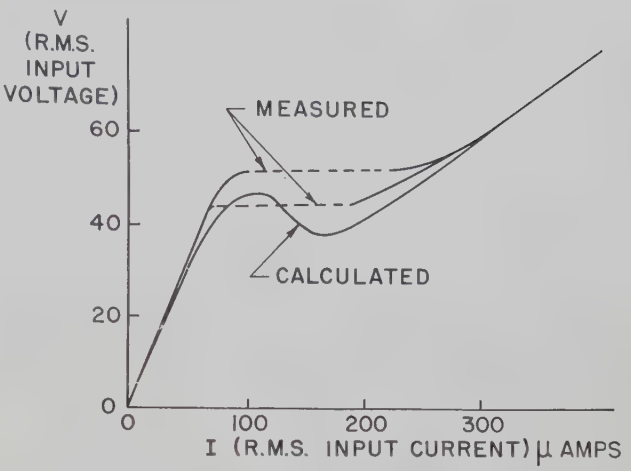


Fig. 6—Steady-state driving-point characteristics of typical regenerative connection.

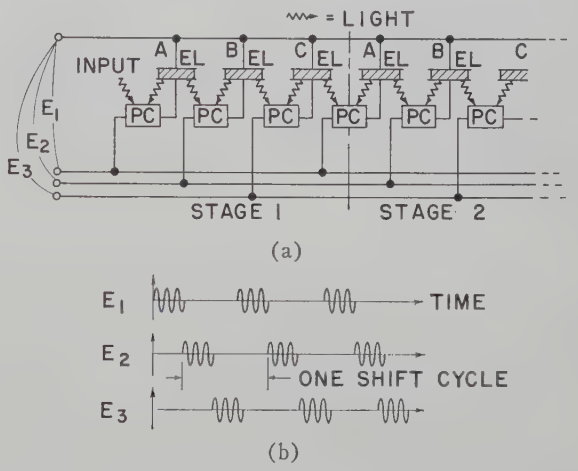


Fig. 7—(a) Electro-optical shift register and (b) shift pulses.

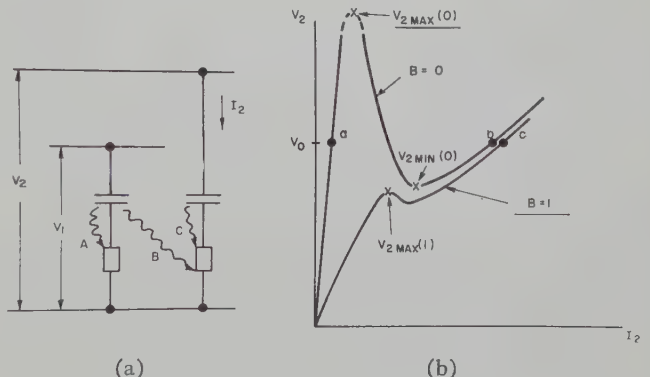


Fig. 8—(a) Coupled elements and (b) V_2-I_2 relation.

Fig. 8(a) shows the electrical and optical connections, while (b) shows the steady-state V_2-I_2 relations. The coupling of the elements is through optical path B. The effect of this radiation upon PC-2 may be considered the same as that of a variable resistance in parallel with the PC in element 2. (An analogous parallel resistance was referred to as R_a in previous discussions.) It may be seen that the V_2-I_2 relation may be altered by the presence or absence of light through path B, referred to as B = 1 or 0.

The operating voltage V_0 must be properly chosen in order that both "1's" and "0's" be shifted. These conditions are shown in Fig. 9. A necessary condition is that $V_{2\min}(0) < V_0$. Additional conditions for "0" and "1" transfer are shown. Since the PC used actually has a time constant, one would design the elements such that the time required to switch above and below the intermediate state shown is about equal.

This configuration was constructed with *ZnS* EL cells and *CdS* PC's, and successfully operated. The excitation frequency was 10 kc; the shifting rate, 30 cps (limited by slow PC's); and the rms voltage , 110 v.

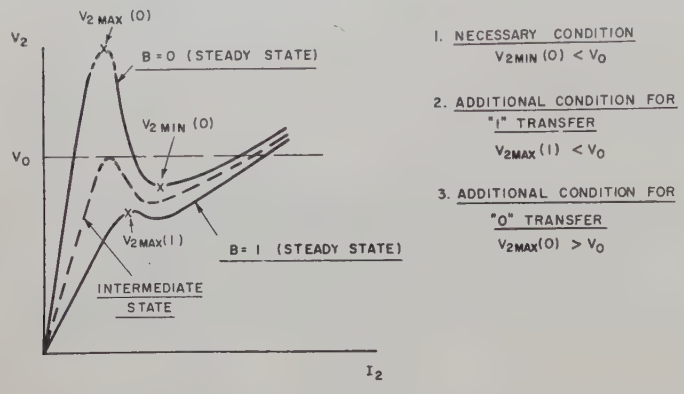


Fig. 9—Conditions for stable "0" and "1" transfer.

Advantages of this shift register are:

1) Inherent indication of state. This, of course, is true only if the radiative spectrum of the EL cell is in the visible region.
2) No other electrical elements needed.
3) Extremely small volume possible. It is feasible with present techniques to produce elements requiring about 10^{-3} inch^3 per bit.
4) Potential low cost—this register does not require an especially accurate time constant in the PC.
5) Information transfer optically with negligible loading effects. Parallel read-out is easily done.

Disadvantages presently known are:

1) Present slow speeds (limited by PC). Speed may be increased by reducing need for sensitivity.
2) Need for ac excitation.
3) Active storage. This is not remediable.

Appendix

Referring to Fig. 1(b), straightforward application of ac circuit theory allows one to find the magnitude of the regenerative connection impedance. Note that phase angle between V and I is not important—only magnitudes are considered. The EL cell voltage is

$$V_{EL} = IX_{CE} \tag{4}$$

where X_{CE} is the capacitive reactance of the EL cell at the carrier frequency f.

$$X_{CE} = \frac{1}{2\pi f C_E}. \tag{5}$$

The total parallel resistance is

$$R_T = \frac{R_{pc}R_a}{R_{pc} + R_a}. \tag{6}$$

The magnitude of impedance is

$$|Z| = (X_{CE}^2 + R_T^2)^{1/2}. \tag{7}$$

Using the power law approximation of (1),

$$R_{PC} = aV_{EL}^{-\eta}. \tag{8}$$

Substituting (8) into (6), the result in (7), and defining

$$\beta \triangleq \frac{X_{CE}}{R_a}, \tag{9}$$

yields:

$$|Z| = X_{CE}[1 + (X_{CE}^{\eta+1}I^\eta + \beta)^{-2}]^{1/2}. \tag{10}$$

The driving-point (rms) voltage-current relation is

$$\begin{aligned} |V| &= |I|\,|Z| \\ &= IX_{CE}[1 + (X_{CE}^{\eta+1}I^\eta + \beta)^{-2}]^{1/2}. \end{aligned} \tag{11}$$

Of interest are the initial and final slopes of the driving-point characteristic, and the valley and peak points. Taking the derivative of (11) with respect to I:

$$\frac{1}{X_{CE}}\frac{\partial V}{\partial I} = \frac{(AI^\eta + \beta)^3 + A(1-\eta)I^\eta + \beta}{(AI^\eta + \beta)^3[1 + (AI^\eta + \beta)^{-2}]^{1/2}}, \tag{12}$$

where

$$A = X_{CE}^{\eta+1}.$$

As $I \to 0$,

$$\begin{aligned} \frac{1}{X_{CE}}\frac{\partial V}{\partial I} &\to (1 + \beta^{-2})^{1/2} \\ \frac{\partial V}{\partial I} &\to (X_{CE}^2 + R_a^2)^{1/2}. \end{aligned} \tag{13}$$

This is the initial slope expected since the photoconductor has very high resistance in the unilluminated state. For large values of current,

$$\begin{aligned} I &\to \infty \\ \frac{1}{X_{CE}}\frac{\partial V}{\partial I} &\to \frac{A^3I^3}{A^3I^3} = 1. \\ \frac{\partial V}{\partial I} &\to X_{CE} \end{aligned} \tag{14}$$

Again, according to the power-law approximation, this is expected. The photoconductors resistance becomes very low (shorting R_a) and the circuit impedance becomes that of the EL cell X_{CE}.

To find the peak and valley points, the derivative is set equal to zero. Thus,

$$(AI^{\eta} + \beta)^3 + AI^{\eta} + \beta - A\eta I^{\eta} = 0. \tag{15}$$

To solve this equation, the following transformation is made:

$$M = AI^{\eta} + \beta$$

$$I = \left(\frac{M - \beta}{A}\right)^{1/\eta}. \tag{16}$$

The equation becomes

$$M^3 + (1 - \eta)M + \eta\beta = 0. \tag{17}$$

Let

$$(1 - \eta) = d \tag{18}$$

$$\eta\beta = f \tag{19}$$

$$M^3 + dM + f = 0. \tag{20}$$

In order that there be two distinct real roots, the following inequality must hold:

$$\frac{\eta^2\beta^2}{4} + \frac{(1 - \eta)^3}{27} < 0. \tag{21}$$

Since $\eta > 1$ for all cases of usefulness,

$$\frac{\eta^2\beta^2}{4} < \frac{(\eta - 1)^3}{27}$$

$$\beta < \frac{2(\eta - 1)^{3/2}}{(27)^{1/2}\eta}. \tag{22}$$

The allowed and forbidden values of the dimensionless ratio β are shown in Fig. 5 as a function of η.

With the condition that the dimensionless ratio β satisfies inequality (22), the roots of (20) can be written in the form

$$M_\delta = -2\left(\frac{-d}{3}\right)^{1/2} \cos\left(\frac{\phi}{3} + 120°k\right), \tag{23}$$

where

$$\delta = 0, 1, 2$$

$$\cos\phi = \left(\frac{-27f^2}{4d^3}\right)^{1/2}. \tag{24}$$

The inequality (22) forces the quantity in brackets in (24) to be positive. Since only positive real values of M_δ are desired, the case $\delta = 0$ can be eliminated, as it leads to a negative M_δ. For $\delta = 1, 2$, two distinct positive real roots are obtained.

If the limiting β for regenerative operation is defined as

$$\beta_0 \triangleq \frac{2(\eta - 1)^{3/2}}{27^{1/2}\eta}, \tag{25}$$

then the roots M_δ can be written as follows:

$$M_1 = \frac{-2}{X_{CE}}\left(\frac{\eta - 1}{3}\right)^{1/2} \cos\left[\frac{\cos^{-1}}{3}\left(\frac{\beta}{\beta_0}\right) + 120°\right] \tag{26}$$

$$= \frac{2}{X_{CE}}\left(\frac{\eta - 1}{3}\right)^{1/2} \cos\left[\frac{\cos^{-1}}{3}\left(\frac{\beta}{\beta_0}\right) - 60°\right]$$

$$M_2 = \frac{2}{X_{CE}}\left(\frac{\eta - 1}{3}\right)^{1/2} \cos\left[\frac{\cos^{-1}}{3}\left(\frac{\beta}{\beta_0}\right) + 60°\right]. \tag{27}$$

After substitution and rearrangement, two quantities may be defined as

$$D_p = 2\left(\frac{\eta - 1}{3}\right)^{1/2} \cos\left[\frac{\cos^{-1}}{3}\left(\frac{\beta}{\beta_0}\right) - 60°\right] \tag{28}$$

$$D_v = 2\left(\frac{\eta - 1}{3}\right)^{1/2} \cos\left[\frac{\cos^{-1}}{3}\left(\frac{\beta}{\beta_0}\right) + 60°\right] \tag{29}$$

The rms peak and valley voltage may be written using (28) and (29).

$$V_{\text{peak}} = (aX_{CE}^{-1})^{1/\eta}(D_p - \beta)^{1/\eta}(1 + D_p^{-2})^{1/2} \tag{30}$$

$$V_{\text{valley}} = (aX_{CE}^{-1})^{1/\eta}(D_v - \beta)^{1/\eta}(1 + D_v^{-2})^{1/2}. \tag{31}$$

The corresponding rms circuit currents are

$$I_{\text{peak}} = X_{CE}^{-1}(aX_{CE}^{-1})^{1/\eta}D_v \tag{32}$$

$$I_{\text{valley}} = X_{CE}^{-1}(AX_{CE}^{-1})^{1/\eta}D_p. \tag{33}$$

Processing Data in Bits and Pieces*

F. P. BROOKS, Jr.†, G. A. BLAAUW†, AND W. BUCHHOLZ†

Summary—A data-handling unit is described which permits binary or decimal arithmetic to be performed on data fields of any length from one to sixty-four bits. Within the field, character structure can be further specified: these processing entities, called bytes, may be from one to eight bits long. Fields may be stored with or without algebraic sign. On all operations, the relative offset or shift between the operand from memory and that from the accumulator can be specified.

Besides the arithmetic operations, three new logical instructions allow any of the sixteen logical connectives of two variables to operate upon each pair of bits in the memory and accumulator operands. The variable field length, variable byte-size features, extend the use of connective operations to a surprisingly wide variety of logical, housekeeping, and editing tasks.

These arithmetic and connective instructions are general and powerful programming tools which greatly simplify complex manipulations. Programming of typical tasks, with both the new instructions and with the instruction set of a conventionally-organized computer, has shown that the new set requires substantially fewer instructions to be written, stored, and executed. Furthermore, the new instruction set has considerably fewer distinct operations than the more conventional set. This is possible because the general-purpose instructions of the new set replace many *ad hoc* instructions which deal with pieces of instructions or data words, or which perform shifting, packing, or editing functions.

The initial application of the variable field length data-processing unit is in the IBM Stretch computer.

Introduction

THE Stretch computer system now under development by IBM has a novel data-handling system. The computer design has been governed by a desire for very high performance and quite general application.[1]

The performance goal was to attain a speed approaching 100 times that of the IBM 704 on typical technical computing problems. This goal was set despite the knowledge that speeds of large core memories could be improved by a factor of only five or six and that circuit speeds could be improved by a factor of only fifty. Improvements in the logical organization of the computer system, therefore, were required to contribute substantially to the increase in performance.

The application goal is somewhat broader than usual. The new computer should perform excellently on large technical and business, computational and logical problems. In the past the design of computers has differed greatly depending on whether they were principally aimed at technical or business applications. Experience has shown that each design is too restrictive, that these two areas do not exhaust the field, and that the computing load of a single installation rarely falls into just one such class.

The breadth of the intended application for the new system forced a re-examination of the way in which data are handled within computers, and of the fundamental units of data in many different kinds of data-processing tasks. This examination led to a new organization which, indeed, contributes considerably to the power and flexibility of the computer.

In addition to a parallel floating point arithmetic unit of very high speed, and another parallel unit for address arithmetic, the new computer has a serial unit for processing data of variable length. This paper, one of a series on the Stretch computer,[2,3] describes the theoretical considerations behind the organization of the variable field length data-handling unit, its logical arrangement and operation, and the functions performed by the new unit in the system as a whole.

Theoretical Considerations

Lengths and Structures of Natural Data Units

In considering automatic data-processing tasks generally, we identify five common types of operations: floating-point operations, fixed-point arithmetic, address arithmetic, logical manipulations, and editing operations. Each of these has a *natural data unit* that is distinct from that of the other types in length, variability of length, or internal structure. An ideal computer would permit each operation to address its natural data unit directly, and this addressing would be simplified by utilizing all properties of the natural data unit that are constant.

It should be observed that the natural data unit is associated with an individual manipulative operation, not with whole programs. In any program, there will be different kinds of operations and therefore different natural data units. Furthermore, the same datum is generally the object of different kinds of operations. For example, a floating-point datum may be developed as a unit in a computation, its components then used in radix conversion arithmetic, and the characters of the result finally used as units in editing for printing. The format of a datum is usually made to agree as closely as possible with the natural data unit of the operations most often performed on that datum.

* Manuscript received by the PGEC, February 26, 1959; revised manuscript received, April 16, 1959. This paper is being presented at the International Conference on Information Processing, Paris, France, June 13–23, 1959.

† IBM Corp., Poughkeepsie, N. Y.

[1] S. W. Dunwell, "Design objectives of the IBM Stretch computer," *1956 Proc. EJCC*, pp. 20–22.

[2] G. A. Blaauw, "Indexing and control-word techniques," *IBM J. Res. & Dev.*, vol. 3; July, 1959. In publication.

[3] F. P. Brooks, Jr., "A program-controlled program interruption system," *Proc. 1957 EJCC*, pp. 128–132.

The natural data unit for most technical computation has come to be the floating-point number, because the use of *floating-point operations* frees the mathematician of many details of magnitude analysis. This unit has considerable internal structure: the representation of a single number includes a number sign, a fraction, an exponent, and, optionally, an exponent sign and bits for flagging numbers (Fig. 1). The fraction part of this unit might be made to vary widely in length, depending upon precision requirements; but the precision analysis which such variation would imply would be as burdensome as the detailed magnitude analysis which floating point operation eliminates. Moreover, these operations must proceed with the utmost speed, and a fixed format facilitates parallel arithmetic. For these reasons floating point numbers follow a rigid format. The datum is usually long; in this machine it uses sixty-four bits, with the fraction occupying forty-eight of these.

Fig. 1—Data unit for floating-point arithmetic.

Fixed-point arithmetic is used on problem data when magnitude analysis is trivial, such as that encountered in business or statistical calculations. Fig. 2 shows some examples. Numbers may or may not be signed. Unsigned binary numbers have the simplest structure as they consist of a homogeneous set of bits. In signed binary numbers, the sign portion usually has properties in which it differs from the rest of the number, thus introducing some additional structure. The numerical part of decimal numbers has a further inner structure of digits which are individually encoded into binary representations. Whether the data unit has a simple structure or a complex one, its natural length is quite variable, with typical numbers varying from four to forty bits in length.

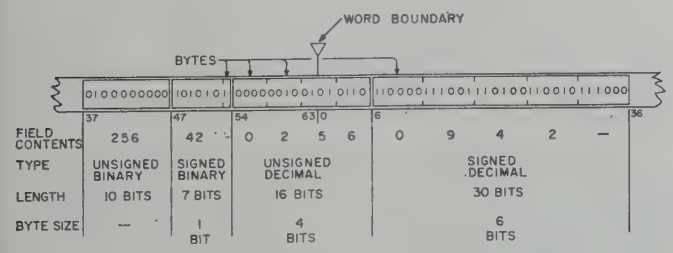


Fig. 2—Data units for fixed-point arithmetic.

Address arithmetic operates upon a natural data unit whose structure is similar to that of fixed-point data, whether decimal or binary (Fig. 3). The unit has, however, one or a few standard lengths because of the fixed size of memory; and the length of the unit is relatively short.

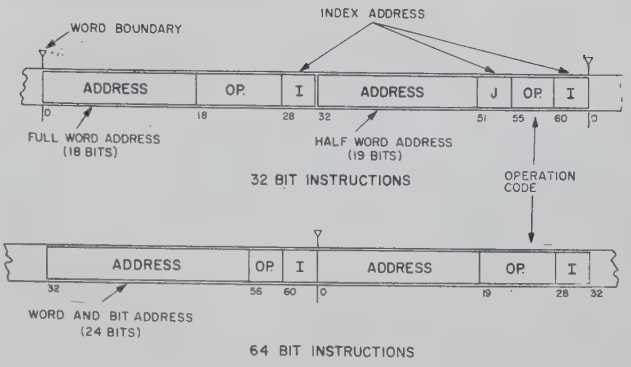


Fig. 3—Data units for address arithmetic.

Pure *logical manipulations*, whether used as a main part of a program, as in combinatorial analysis, or used for controlling the course of the program, operate upon very simple data units. A data unit consists of a group of bits each of which has an independent meaning (Fig. 4). This distinguishes such operations from arithmetic, which uses bits as components of numbers. Since field length in logical operations depends upon the number of operations that can be paralleled, lengths vary; but since no carries are propagated, restriction to fields of arbitrary lengths is not too burdensome.

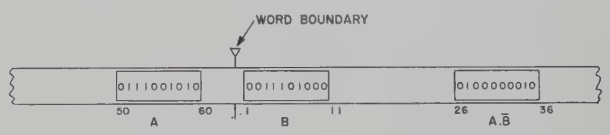


Fig. 4—Data units for logical manipulations.

A final class, *editing operations*, includes all operations in which data are transformed from one format to another, checked for consistency with a source format, or tested for controlling the course of the program. The natural data unit for such operations varies widely (Fig. 5). All the natural data units of the previous four types of operation undergo editing operations in the normal course of their processing; and there are other units that are unique to editing operations. For example, a group of data fields may be moved as a unit within memory to assemble records for output.

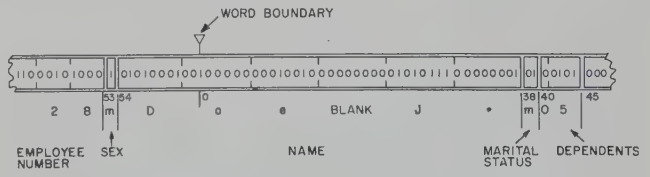


Fig. 5—Data units for editing operations.

Editing operations not only possess the most complex natural data structures, but they also use the most widely varying natural field lengths. For some manipulations the natural unit is the single character; for other manipulations, such as comparison or transmission, the natural data unit is a field of many characters.

Besides these five kinds of natural data units that can be identified for operations commonly built into computers, other natural data units are suitable for operations that are usually encoded with subroutines, such as matrix arithmetic, complex arithmetic, and multiple precision arithmetic. These larger units need not be considered separately, as they are necessarily composed of components that are themselves the data units of some built-in operation.

Possible Procedures for Handling Natural Data Units

The previous section has shown how natural data units for different operations differ in structure, length, and variability of length. Since a general-purpose computer must deal with all these units, its organization must accommodate their diversities.

The data and instructions for any given problem can be considered to consist of a single stream of natural data units, without computer-prescribed spacers, groupings, etc. The computer designer must furnish a memory structure and an addressing system with which the individual components of a stream of natural data units can be conveniently manipulated. Subsequently, the problem programmer must map the problem-data stream into a spaced and grouped stream suitable for the memory organization which the computer designer provides. This mapping requires some of the computer's power and necessarily introduces some inefficiencies. The more complex and difficult the mapping, the lower is the performance of the whole system.

The diversity of the natural data units implies that more information is required for their specification than if they were alike. The computer designer can choose the manner in which the user pays this information price, but the price must be paid.

The classical approach to this problem was to ignore it. For simplicity, early computer designers assumed that provisions for handling the object data of fixed-point arithmetic operations would suffice, and that the natural data unit for these operations was the single number of constant length. These two assumptions led to a simple, homogeneous, fixed-word-length memory organization. Since neither assumption was completely true, the information price of diversity was paid by the user in reduced performance and more complex programming.

When performing operations other than fixed-point arithmetic, such as editing and address arithmetic, the programmer shifted, extracted, and packed, in order to get at the natural data unit of the operation. When faced with data of varying lengths, the programmer had two options as to the method of paying the information price. He could place each unit in a different machine word, or he could pack several of the shorter units into a single word. (Since the machine word was usually picked to be a reasonable upper bound on natural data lengths, he was less often faced with the problem of manipulating units that required several words.) If he chose to use a different addressable word for each data unit, he paid the information price in reduced memory capacity and in longer operating times for input-output and arithmetic units. The programmer's other option was to provide a number of so-called housekeeping instructions, mostly shifts, for packing and unpacking addressable memory cells. The price is paid in memory capacity for the extra instructions, in execution time, and in programming time.

Clearly, one way to improve the performance of a computer by changing its organization is to pay the price of diverse data units in the form of more complex hardware rather than reduced performance. This implies a memory structure that can be composed of variable length cells, and several computers have been so organized.[4] These computers have been intended primarily for business data processing, where editing operations are of great importance and where the assumption of constant-length data units is particularly poor. As the importance of nonarithmetic operations in all kinds of calculations become more apparent, the conclusion crystallized that a variable cell length memory organization is essential for any high-performance, general-purpose computer.

There are several methods of achieving variable cell size. If the memory is to be addressed rather than scanned, the cell lengths may vary from cell to cell and from problem to problem; but the positions (and therefore the lengths) of cells must remain constant within a single computation. That is, cells at different addresses may have different lengths, but a change of contents of a cell must not change its length. On tape, where scanning is used instead of addressing, this constraint does not hold, and some computers allow item lengths on tape to vary by deleting leading numeric zeros or trailing alphabetic blanks.

A simple way of organizing a memory of different cell sizes is to provide a fixed complement of assorted sizes. This rather inflexible arrangement was discarded in favor of a second method where the smallest component of a memory is made addressable; a cell is defined by specifying both the position of one component and the extent of the cell. Because of the requirements of pure logical operations and of editing operations, addressing resolution was provided all the way down to the individual bit level. Each bit of the memory has a unique address.

There are several techniques of specifying cell extent. The first is to use a unique combination of bits as a partition between cells. This method is used to separate numerical fields in the IBM 705. The use of partition symbols implies reduced memory capacity due to the symbols themselves and, more seriously, exclusion of the partition bit combination from the set of permissible data symbols. This difficulty alone would have precluded use of partitions between memory cells in the Stretch computer, because arbitrary bit combinations

[4] For example, C. J. Bashe, W. Buchholz, and N. Rochester, "The IBM Type 702, an electronic data processsing machine for business," *J. ACM*, vol. 1, pp. 149–169; October, 1954.

arise in assembling instructions, reading data from external devices, and performing binary computations. Such activities could not be excluded. Furthermore, in any computer where memory speed is the limiting factor on performance, it is highly desirable that each bit fetched from memory contain one bit of information. Use of extra symbols and restrictions on bit combinations both reduce information content.

A second method of specifying cell extent is to use a Procrustean Bed technique in which data are transferred from memory to a register until the register is full. Transfers to memory likewise proceed until the register is completely copied. This technique is used for alphabetic fields in the IBM 705. The disadvantage is that the technique requires many extra instructions for changing the length of the receiving register, or the use of several receiving registers of different lengths.

A third technique, and that adopted, is to provide the information on cell extent in the instructions which use that cell. This can be done by specifying one of several masks, by specifying beginning and end, or by specifying beginning and length. In order to simplify indexing, the last method was selected. Each instruction that can refer to variable-length cells contains the complete address of the leftmost (high order) bit of the cell and the length of the cell; however, instructions that do not need to refer to cells of varying length are not burdened with all this information.

The Variable Field Length System

Addressing

To obtain a high data rate from core memories, it is necessary to define memory words with a large number of bits that are read or written in parallel. The memory word length was set at 64 information bits, plus error-detection and -correction bits. This choice, a power of two, permits word addresses and bit addresses to be manipulated and indexed with a uniform radix arithmetic. Carries from binary operations on bit addresses meaningfully propagate into word addresses.[2]

To be practical, floating-point number representations must be longer than 32 bits, but 64 bits are ample. With 64-bit floating-point numbers, multiple precision operations can usually be avoided; this radically improves performance on some problems.

Sixty-four bits is also a convenient word size for instructions that must specify the several parameters of a variable-length field. Instructions operating upon the more rigid formats of floating-point and address arithmetic do not need all the variable-field specifications, and these instructions were abbreviated to 32 bits.[5] Providing both full-length and half-length instructions lead to more efficient use of instruction bits and substantially improves the over-all computer performance. A full-length instruction, incidentally, is not restricted to a single memory word; it may occupy adjacent halves of two consecutive words. Thus, no gaps are left when full- and half-length instructions are intermixed.

The maximum size of directly addressable memory was chosen as 2^{18} (262,144) words. Addressing to the bit level then requires 24 bits, of which the 18 high-order bits specify the memory word and the 6 low-order bits refer to one of the 2^6 (64) bits within that word. For addresses referring only to full words, such as are needed in floating-point operations, 18 bits are sufficient. To address half-words, 19 bits are used; thus, instruction locations are 19 bit addresses.

The high-order 18 bits of each kind of address always refer to a full-word location, and any additional bits specifying subdivisions of a word are placed to the right of the basic 18 bits. For simplicity, the left ends of all addresses were aligned within the format. The obvious way to do this was to put the address on the left of the instruction as shown in Fig. 3. All formats are variations on the basic pattern: Address, Operation, Index. It is found to be very desirable to make all addresses fully indexable, and, with one exception, this has been achieved.

For operations upon fields of variable length, it is generally necessary to specify the inner structure of the field. For alphabetic fields this consists of the individual letters or other characters. For numeric fields, the structure includes the sign, if any, and the digits, if separately encoded. These sub-units collectively have been named *bytes*. Since the coded representation of a byte naturally varies in size, byte sizes of one to eight bits may be specified and used as illustrated in Fig. 2.

Another specification in the variable field length instruction is the *offset* or relative shift, to be used when combining the cell contents with the accumulator contents. This is very similar to the shift specification on the IBM 604, and it eliminates the need for separate shifting operations and instructions.

In all variable-field-length operations, the word and bit address itself can be used as the data rather than the address of the data. This *immediate addressing* is specified in the format.

The Mechanism

The variable-field-length mechanism is shown in a highly schematic form in Fig. 6. It includes an accumulator, two words (128 bits) long, with a separate register for holding sign and flag information. Associated with the accumulator is a two-word nonaddressable receiving register, *CD*. All data transfer takes place in parallel, 64 bits at a time, between this register and memory. Since field lengths are restricted to 64 bits, any field is completely contained in not more than two consecutive memory words. An operand is fetched from memory by bringing the words containing it to the receiving register. Memory modification also requires the words containing the addressed field to be fetched, so that bits adjacent to the addressed field may be preserved.

The arithmetic and logical units handle up to eight bits in parallel and operate serially by byte. A transistor

[5] W. Buchholz, "The selection of an instruction language," *1957 Proc. WJCC*, pp. 128–130.

switch matrix selects the bytes of a field one after another, beginning at the right end. Each byte is masked to eliminate unwanted bits, under control of the specified byte size. The masked byte passes into the adder or logical unit, where it is combined with a similarly selected byte from the accumulator. (A more detailed description of this rather complex mechanism is expected to be given in a subsequent paper.)

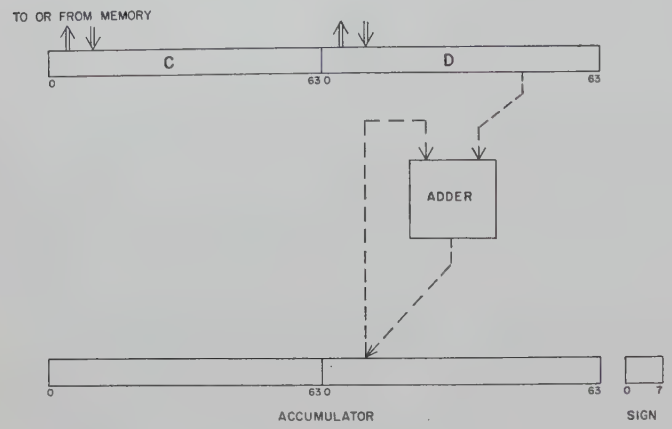


Fig. 6—Schematic of variable-field-length mechanism.

Arithmetic Operations and Features

The variable-field-length system provides decimal and binary addition and comparison operations, direct binary multiplication and division, and automatic decimal-binary and binary-decimal conversion.

The basic arithmetic operations performed by the variable-field-length mechanism are LOAD, STORE, ADD, ADD TO MEMORY, and COMPARE. Each of these has variations and modifications. For example, a variation of ADD, called AUGMENT, permits the accumulator sign to be ignored and causes the accumulator contents to be set to zero if the result attempts to change sign.[6] In multiplication, division, and conversion operations, the variable-field-length system is used to transform and place the operands suitably for performing the actual operation in the parallel floating-point unit. The result is then aligned as specified. This procedure permits these operations to utilize both the flexibility of the variable-field-length mechanism and the speed of the floating-point mechanism. The alignment procedure removes common restrictions upon division: any divisor other than zero yields a valid quotient.

It proved to be faster to perform decimal multiplication and division by a subroutine performing conversions and parallel binary operations than by built-in serial decimal operations. The decimal MULTIPLY and DIVIDE operation codes give automatic operand aligning and subroutine entry functions.

Since arithmetic may operate upon quantities whose signs are either expressed or implied, the operation code may specify whether the field contains a sign or not. If a sign is present, it is contained in a separate sign byte that contains flag bits as well. The sign byte is of the specified size. (This is the only function of the byte-size specification in binary arithmetic.)

A sign modifier bit in each instruction permits the sign of one of the operands to be considered as inverted for the operation. This not only permits additions to be modified into subtractions, it also permits sign inversions on loading, storing, and other operations.

The accumulator byte size is fixed at four bits for decimal arithmetic. Whenever a decimal quantity is loaded into or added to the accumulator, its bytes are cut or expanded to four-bit size. Conversely, when the accumulator contents are stored, a similar contraction or expansion takes place.

Whenever an operand is fetched from memory, its flag bits are sensed and program interruption indicators are actuated.[3] When an operand is stored, flag bits may be attached. Other interruption indicators are actuated by exceptional conditions arising during arithmetic operation. Examples are attempts to load data or propagate carries beyond the left end of the accumulator, or to ignore significant high-order bits when storing accumulator contents. Still other indicators may be set if invalid or ill-defined operations are attempted.

The result of each arithmetic operation is reflected in the setting of certain result indicators. These indicators show whether the result was greater than zero, equal to zero, or less than zero; whether the result was placed in memory or not; or whether a comparison showed the accumulator to be low, equal, or high.

Appendix I shows some short examples of the use of the variable-field-length arithmetic operations.

Logical Operations and Features

A logical operation called CONNECT provides any of the sixteen logical connectives of two variables, such as *and*, *or*, *not*. CONNECT is a variable-field-length type of instruction which specifies a memory field to be combined, bit by bit, with the accumulator field according to a specified connective. The result is returned to the accumulator or to memory, or it may be discarded. In any case, indicators are set and counts are developed.

The connective is specified by four bits in the instruction. Following directly from this specification, the logical unit for each bit consists simply of four three-way *and* circuits feeding a four-way *or* circuit. Each *and* circuit combines one of the connective specifier bits from the instruction with the two operand bits or their appropriate inverses. Indicators are set according to whether the result is all zeros. A count is developed of the total number of ones in the result. A second count is

[6] The augment operation is a modification of operations independently proposed by:

F. P. Brooks, Jr., "The Analytic Design of Automatic Data Processing Systems," Ph.D. thesis, Harvard University, Cambridge, Mass., sec. 6.42; 1956.

R. W. Murphy, "A positive-integer arithmetic for data processing," *IBM J. Res. & Dev.*, vol. 1, pp. 158–170; April, 1957.

developed of the number of zero bits to the left of the leftmost one bit in the result.

Since logical operations have a natural data unit with extremely simple structure, byte-size control seems superfluous. Because the byte-size mechanism was available, it was decided to use it to provide for expansion, interlacing, and contraction of fields. The accumulator operand for all connective operations is considered to be composed of eight-bit bytes. The memory field is considered to be composed of bytes of the specified size. During a logical operation, the right end of each of these bytes is aligned with the right end of the corresponding eight-bit accumulator byte. The high-order bits of the bytes from memory are assumed to be zero. If a memory byte size of eight is specified, no byte-size transformation is performed. The memory field length is not required to be a multiple of the specified byte size.

It should be remarked that the provision of all sixteen connectives was originally proposed for the sake of completeness and as a matter of principle, though it appeared that *not*, *and*, *or*, and *exclusive or* would in practice suffice. As the format was defined and such a simple circuit resulted for mechanizing the connectives, it became apparent that the extra cost of furnishing the other connectives was very low indeed. Exploratory programming revealed the surprising fact that the connectives involving both the memory bit m and the accumulator bit a were used much more rarely than what would appear to be the "trivial" connectives (m, $\bar{m}$, a, $\bar{a}$, 0 and 1). The reason is that these connectives furnish useful techniques for setting, resetting, inverting, and testing bits in the accumulator or in memory and for loading and storing parts of the accumulator independently of other parts.

Appendix II contains an example of the usefulness of the new logical operations in a matrix multiplication problem which is ordinarily considered to be purely arithmetic computing.

Function of the Variable-Field-Length Operations in the Over-all System

The variable-field-length operations of the IBM Stretch computer play several important roles in improving over-all computer performance. In some areas, they permit novel computational techniques. In others, they facilitate well-known but little-used techniques. In still other areas, they furnish facility and economy in accomplishing common tasks in a straightforward manner.

Fixed-Point Computing

In the area of fixed-point computing, the variable-field-length arithmetic system does indeed furnish the efficiencies that one would expect from theoretical considerations. Memory can be used economically for storage of data. Operating times of serial input-output and arithmetic units are reduced because fields have only the required lengths.

Since almost all programs involve some fixed-point computing as part of housekeeping procedures, these efficiencies contribute to performance improvement in all types of problems. Even in technical calculations where most instructions are floating-point arithmetic or address arithmetic, the new computer typically requires fewer instructions written, fewer instructions stored, and fewer instructions executed, than do present fixed-word-length technical computers, such as the IBM 7090 and its predecessors.

Instruction Set Simplification

Early computers had quite simple instruction sets. As experience was gained, the value of added functions was recognized and many new operations were added to the old instruction sets. A good number of these were operations addressing parts of words, *i.e.*, addresses, index tags, exponents, fractions, decrements, signs, etc. In the IBM 7090, approximately one-quarter of the operations are of this sort. In the Stretch computer, the general-purpose variable-field-length operations substitute for many of these *ad hoc* operations. This simplification makes the instruction set for the new computer easier to learn and to remember.

Table Reference

The variable-field-length system provides a general-purpose mechanism for table references by either direct-reference or search techniques. As memories grow larger, these techniques become increasingly useful as substitutes for involved algorithms. The variable-field-length system facilitates the use of these techniques in several ways. Tables can be compressed to a minimum. Any bit combination can be a table argument. Finally, the uniform radix of the addressing system offers the simplest possible transformation from data to address.

Logical Operations

The facilities for logical operations constitute perhaps the most important new feature of the computer. It is now clear that such operations are neither modifications of arithmetic nor auxiliaries to it, but are equal to arithmetic in importance. The logical facilities, in conjunction with the variable-field-length mechanism, permit transformation of data formats (selection, expansion, contraction, interleaving, and distribution of bits), transformation of data contents (setting, resetting, and inverting bits and connecting pairs of bits), and analysis of the resulting data (left-zeros and all-ones counts and other indications). The total constitutes a complete, novel, and powerful system for operating upon groups of independent bits rather than numbers.

Editing Operations

The facilities for loading, storing, and comparing variable-length fields and for performing binary logic constitute a general and powerful editing system. They simplify format conversion, tests for data consistency, and tests to guide the course of calculation. This editing

power is especially useful in dealing with constrained data formats of devices such as printers, plotters, and punched-card units. Special operations for particular tasks were avoided by providing general-purpose facilities which could readily be used to solve a wide variety of editing problems.

Conclusion

The natural bits and pieces of data can be processed in a straightforward manner by the variable-field-length mechanism described. Such a mechanism extends the breadth of application of any computer and improves performance on almost all kinds of tasks. The high goals set for the IBM Stretch computer demanded improvements in organization. Exploratory programming has already confirmed the belief that the variable-field-length system constitutes a substantial improvement.

Appendix I

Examples of Variable-Field-Length Arithmetic

Example 1

The third field of Fig. 2 is to be added to the fourth field with the result remaining in the accumulator. The third field is assumed to have its decimal point between the third and fourth digits from the left, and the fourth field has its point between the second and third digits. The words shown are assumed to have (decimal) addresses 1789 and 1790.

Operation	*Modifiers*	*Word and Bit Address*	*Field Length*	*Byte Size*	*Offset*
LOAD	Decimal, Unsigned	1789.54	16 bits	4 bits	4 bits
ADD	Decimal, Signed	1790.06	30 bits	6 bits	0
	First operand		025.6		
	Second operand		09.42−		
	Result		016.18+		

Example 2

The second field in Fig. 2 is to be doubled and added to the first field, with the result replacing the first field in memory. The result is expected to remain positive.

Operation	*Modifiers*	*Word and Bit Address*	*Field Length*	*Byte Size*	*Offset*
LOAD	Binary, Signed	1789.47	7 bits	1 bit	1 bit
ADD TO MEMORY	Binary, Unsigned	1789.37	10 bits	—	0

	Binary	*Decimal Equivalent*
First operand	101010 −	84−
Second operand	0100000000	256
Result in Memory	0010101100	172

The use of offset 1 in the first instruction effectively multiplies the number loaded by 2^1. Had the result been negative, an indicator would have been actuated to allow automatic program interruption. The unsigned modifier on the second instruction refers to the memory operand; the negative accumulator sign affects the operation in the usual way.

Appendix II

Example of the Use of Logical Facilities in the Solution of an Arithmetic Problem

Two matrices, consisting mostly of zero elements, are to be multiplied. Each matrix is represented by a matrix of single bits, each bit corresponding to an element. One bits represent nonzero elements. In addition, the values of all nonzero elements are listed in memory in consecutive row or column order. Thus, storage space has been collapsed by omitting all zero elements.

In forming the vector inner products, only nonzero elements need to be multiplied. Pairs of nonzero elements in two vectors are found by *and*ing the corresponding bit vectors and looking for one bits in the result. The *left zeros count* identifies the position of the first pair. This count is used as a *field length* for two subsequent logical operations, each of which examines one of the original bit vectors. Each of the resulting *all-ones counts* gives the number of nonzero elements preceding that desired, and, therefore, the position of the desired element in the compacted list. Using indexing techniques to locate the operands, their product is formed and added to the partial sum. The procedure is iterated through each pair of vectors and from vector to vector.

Such bit matrix techniques can, of course, be used in any binary computer to save considerable time by omitting lengthy multiplications. An actual program comparison, however, reveals that much more is to be gained from the extensive logical facilities of the Stretch computer. The comparison is with the most advanced of IBM's series of technical computers of conventional design, the IBM 7090. The program for each machine was written to utilize techniques most efficient for that machine. The matrices were assumed to be of order 36 with ten per cent nonzero elements.

	7090	*Stretch computer*	*Ratio*
Instructions executed	242,600	23,160	10.5
Instructions written	73	41	1.8
Program storage (words)	73	29.5	2.5
Program storage (bits)	2628	1888	1.4

Note that the faster operating speed of the Stretch computer does not affect this comparison. A further gain is obtained in going to higher orders because of a longer floating-point word length and better double-precision facilities.

Although the logical facilities were designed to be general, they turned out to be especially apt for this problem. The increase in system performance due to logical power is not always so striking, but it appears to some extent in almost all problems.

Acknowledgment

Substantial contributions to the planning of the variable-field-length system outlined in this paper were made by S. W. Dunwell, C. A. Scalzi, and W. Wolensky.

Increasing Reliability by the Use of Redundant Machines*

D. E. ROSENHEIM† AND R. B. ASH‡

Summary—The improvement of reliability and availability through redundancy of entire machines rather than of components is investigated. An attempt is made to break down the cost of operating a digital computer, and to determine the relationship between cost and system failure. Three specific cases are discussed.

Case 1: Where n machines are operated independently, processing the same input data. The output is taken from a single one of them; if this machine fails, the output is promptly switched to a machine which is operating properly. As soon as repairs can be completed, the machine which had failed is returned to operation. System failure occurs only when all n machines are in the failed condition at the same time. A penalty cost is assessed for system failure, this cost being proportional to the system down-time.

Case 2: Where n machines are operated as in Case 1, except that any machines which fail are not returned to operation until the beginning of the next operating period. Penalty cost for system failure is assessed in the same way as in Case 1.

Case 3: Where n machines are operated as in Case 2, but where the penalty cost for system failure is a fixed amount and is independent of the resulting down-time.

Introduction

IT HAS been stated[1] that the fabrication industries have lagged behind the process-control industries in the use of automation techniques. The lack of precision of analog computing techniques is certainly one of the limitations of the use of analog-computer control in large fabrication industries. While the use of digital-computer control of plant operation would overcome the lack of precision in the analog machines, there is still much hesitation by management to entrust the operation of plants to large digital computers. For these applications the reliability of present digital computers is often not adequate. However, means are available to increase computer reliability so that the field of application of such computers for use in on-line operations may be extended. In this paper we define reliability as: "The probability of adequate performance of a specified function, for a given period, under specified conditions."

Using presently available components, the reliability of digital computers may be increased by two general methods. The first of these is an adequate preventive maintenance program with marginal checking to anticipate the large majority of component failures before they have a chance to cause machine trouble.

Although much theoretical work has been done on improving reliability through the use of redundant circuitry,[2–5] very little has been done from a practical standpoint to show how to use such redundant circuits, relating the increased reliability to the cost of redundancy. B. J. Flehinger[5] has shown how the reliability improvement depends upon the machine level at which redundancy is applied, and concludes that for initially reliable machines the increase in reliability is substantially obtained from the degree of redundancy and is only secondarily affected by the level at which the redundancy is applied. Redundancy of complete machines allows for the use of already existing machines to obtain a very reliable system; redundancy at the small-unit level requires that a special-purpose system be built for each desired over-all system reliability. Furthermore, small-unit redundancy requires the extensive use of switching elements to choose between good or failed units. In addition, the use of redundant units may cause a serious maintenance problem in finding and replacing failed components of redundant units. For these reasons, we feel that the practical advantages of redundancy of complete machines for applications where a failure is costly justifies investigation of the relationship between increased reliability and the cost of redundancy.

In order to give an idea of the possible advantages of redundant machines, let us attempt to evaluate the cost of a digital computer for a potential customer. This cost may in general be divided into two major categories.

1. *Normal Operating Cost*

This includes rental and maintenance, power consumption, and air conditioning. For more than one machine, therefore, the normal operating cost should be proportional to the number of machines.

2. *Penalty Cost (Due to System Failure)*

We define penalty cost as the cost per unit operating time resulting from system failure. (If the system application permits, the machines may be programmed to recompute a portion of the program whenever an error occurs. In this case, only repeated errors would be con-

* Manuscript received by the PGEC, February 3, 1959.

† IBM Watson Lab., Columbia University, New York, N. Y.
‡ Dept. of Electrical Engineering, Columbia University, New York, N. Y. Formerly IBM Watson Lab., Columbia University.

[1] E. M. Grabbe, "Automation in Business and Industry," John Wiley and Sons, New York, N. Y.; 1957.

[2] J. von Neumann, "Probalistic Logics," Automata Studies, No 34, Princeton University Press, Princeton, N. J.; 1956.

[3] C. E. Shannon and E. F. Moore, "Reliable circuits using less reliable relays," *J. Franklin Inst.*, September–October, 1956.

[4] W. E. Dickinson and R. M. Walker, "Reliability improvement by the use of multiple-element switching circuits," *IBM J. Res. and Dev.*, vol. 2, pp. 142–147; April, 1958.

[5] B. J. Flehinger, "Reliability improvement through redundancy at various system levels," *IBM J. Res. and Dev.*, vol. 2, pp. 148–158; April, 1958.

sidered as failures. In some applications this will not be possible and it will be necessary to consider any error as a machine failure.) The penalty cost depends upon the assumed operating conditions and will be specified for each of the three cases discussed below.

We may now begin the discussion of machine redundancy by considering the first of three modes of operation.

Case 1

Our assumptions are:

1) There are n machines operated independently, processing the same input data. The output is taken from a single one of them; if this machine fails, the output is promptly switched to a machine which is operating properly.

2) As soon as a machine is repaired, it is immediately returned to operation.

3) The time required to make a repair is independent of the time of occurrence of the failure. All repairs are made during scheduled operating time.

4) The preventive maintenance time is not part of the scheduled machine operating period.

5) System failure occurs only when all n machines are in the failed condition at the same time.

6) The penalty cost is proportional to the average system down-time per unit of scheduled operating time.

7) The failure density function for each machine is $f(t) = (1/T_m) \exp(-t/T_m)$, where t is time measured from the beginning of the working day and T_m is the mean time to failure. This is the exponential failure law, and has been justified for a wide class of complex systems for periods of observation small compared with the mean time to failure of an individual component.[6]

The average cost per unit time for operating one digital computer may therefore be written as

$$C_1 = K + P\delta$$

where

K = the normal operating cost per unit of scheduled operating time,

δ = the average down-time per unit of scheduled operating time (with the above assumptions this will be equal to the probability that the machine is inoperative at any given time during scheduled machine operation),

P = the extra cost per unit of down-time.

Given any particular machine, δ can be determined from adequate field records of machines in use. It is for the potential user to estimate, for his particular application, the value of P, or, in other words, the cost damage to his plant operation for a unit of down-time of the computing system.

[6] D. R. Cox and W. L. Smith, "On the superposition of renewal processes," *Biometrika*, vol. 41, pp. 91–99; 1954.

For two identical machines running independently, the cost per unit time is given by

$$C_2 = 2K + P\delta^2,$$

since δ^2 is the probability that at any given time during scheduled operation both machines are inoperative. Similarly, for n identical machines the cost per unit time is given by

$$C_n = nK + P\delta^n.$$

The integral value of n which will minimize the overall computer cost C_n is determined as follows.

Let

$$C(v) = vK + P\delta^v,$$

where v can assume any real positive value.

That integer, n, which minimizes C_n, lies between two values v and $v-1$, for which

$$C(v) - C(v-1) = 0.$$

This equation yields

$$v = 1 + \frac{\ln\left[\dfrac{K}{P(1-\delta)}\right]}{\ln \delta},$$

and n is the greatest integer less than v.

We have avoided mention of the considerations involved in operating multiple machines into a common output. We have also neglected the necessary switching of machines and transferring of information from one machine to another when a machine is taken out of service to be repaired or brought back into service after having been repaired.

It is a relatively simple matter to remove a single machine from the system if it fails. However, after this machine is repaired it is necessary to return it to operation. This necessitates the transfer of the information in the operating machines into the repaired machine, and therefore requires additional equipment for a fast transfer, and also requires that the controlled operation be temporarily halted for the information transfer. This transfer need only take a fraction of a minute, and so in many operations the plant stoppage for this amount of time could be tolerated. However, in some applications the construction of complicated equipment for information transfer may not be desirable. A new mode of operation results when no transfer of information between machines is allowed.

Case 2

Our assumptions are:

1) There are n machines operated independently, processing the same input data. The output is taken from a single one of them; if this machine fails, the output is promptly switched to a machine which is operating properly. Once a machine has failed it remains inoperative for the remainder of the working day.

2) All machines run maintenance-free until a time T_0, which corresponds to the end of the working day. All machines are in perfect working order at the beginning of the next day.

3) System failure occurs only when all n machines are in the failed condition at the same time.

4) The failure density function for each machine is $f(t) = (1/T_m)\exp(-t/T_m)$, where t is time measured from the beginning of the working day and T_m is the mean time to failure. This is the exponential failure law.

5) The penalty cost is proportional to the average system down-time per unit of scheduled operating time.

A Single Machine

Once again the cost per unit time of operating a single machine is given by

$$C_1 = K + P\delta_1.$$

If the machine fails at a time t, the down-time is $T_0 - t$, since the machine is out of operation for the remainder of the working day. The down-time per unit of scheduled operating time is $(T_0 - t)/T_0$. The mean value of this quantity is δ_1 and is given by

$$\delta_1 = \int_0^{T_0} f(t)\,\frac{(T_0 - t)}{T_0}\,dt.$$

Substituting the assumed density function, we find

$$\begin{aligned}\delta_1 &= \frac{1}{T_m}\int_0^{T_0}\frac{(T_0 - t)}{T_0}\exp(-t/T_m)dt \\ &= 1 - \alpha(1 - e^{-1/\alpha})\end{aligned}$$

where

$$\alpha \equiv T_m/T_0.$$

The total cost is:

$$C_1 = K + P[1 - \alpha(1 - e^{-1/\alpha})].$$

n Machines in Parallel

The cost of operating n machines is

$$C_n = nK + P\delta_n,$$

where δ_n is the mean system down-time per unit of operating time.

$$\delta_n = \int_0^{T_0}\frac{T_0 - \tau}{T_0}f_n(\tau)d\tau.$$

where τ is the time of system failure (*i.e.* when the nth machine fails). But

$$f_n(\tau)d\tau = dF_n(\tau)$$

where $F_n(\tau)$ = the cumulative failure distribution function for the system consisting of n identical independent machines. Since $F(t) = (1 - e^{-t/T_m})$ for each machine,

$$F_n(\tau) = (1 - e^{-\tau/T_m})^n,$$

and

$$\delta_n = \int_0^{T_0}\frac{T_0 - \tau}{T_0}\,d(1 - e^{-\tau/T_m})^n.$$

Integrating by parts this yields

$$\delta_n = \frac{1}{T_0}\int_0^{T_0}(1 - e^{-\tau/T_m})^n d\tau = \alpha\int_0^{1/\alpha}(1 - e^{-s})^n ds.$$

On making use of the binomial theorem to expand $(1 - e^{-s})^n$ and integrating the series term by term, we find

$$\begin{aligned}\delta_n = 1 - \alpha\Bigg[&n(1 - e^{-1/\alpha}) - \frac{n(n-1)}{2!\cdot 2}(1 - e^{-2/\alpha}) \\ &+ \frac{n(n-1)(n-2)}{3!\cdot 3}(1 - e^{-3/\alpha}) - \cdots \\ &+ (-1)^{n-1}\frac{1}{n}(1 - e^{-n/\alpha})\Bigg].\end{aligned}$$

As for Case 1, we define a function

$$C(v) = vK + P\delta(v)$$

where v is any positive real number, and find the integer n between $v - 1$ and v where

$$C(v) - C(v - 1) = 0.$$

Now

$$C(v) - C(v - 1) = K + P[\delta(v) - \delta(v - 1)],$$

and

$$\begin{aligned}\delta(v) &= \alpha\int_0^{1/\alpha}(1 - e^{-s})^v ds \\ &= \alpha\int_0^{1/\alpha}[(1 - e^{-s})^{v-1} - e^{-s}(1 - e^{-s})^{v-1}]ds \\ &= \delta(v - 1) - \frac{\alpha}{v}(1 - e^{-1/\alpha})^v.\end{aligned}$$

Therefore

$$C(v) - C(v - 1) = k - \frac{P\alpha}{v}(1 - e^{-1/\alpha})^v.$$

On setting this last expression equal to zero, we find

$$v = \frac{\alpha P}{K}(1 - e^{-1/\alpha})^v.$$

We may write this result as

$$v = P\alpha/Ku,$$

where

$$u \ln u = (P\alpha/K)\ln\{1/(1 - e^{-1/\alpha})\}.$$

Since the right-hand side is a positive constant, there is a unique value of u greater than one which is the root of this equation. Correspondingly, there is a unique n, the greatest integer *less than* v, which minimizes the cost C_n. Clearly,

$$n < \frac{P\alpha}{K}.$$

We may define an improvement factor for n-machine operation as

$$I_n \equiv \frac{\delta_1}{\delta_n} = \frac{\text{mean down-time for 1-machine operation}}{\text{mean down-time for } n\text{-machine operation}}.$$

Fig. 1 shows a plot of I_n α for $n=2$, 3, and 4. Fig. 2 shows a plot of I_n vs n for three values of α. It is well to note that for quantitative work the analytic expressions should be used for it is difficult to read the graphs accurately.

For a maintenance-free interval which is small with respect to the mean time to failure, ($\alpha \gg 1$), the improvement factor is large. However, if the penalty cost $P\delta_1$ is already small for one-machine operation, the addition of another machine will not be indicated, since normal operating cost will be the decisive factor. In other words, we must consider the absolute reduction in average down-time. For $\alpha=4$, $\delta_1=0.112$ for one-machine operation, and $\delta_2=0.012$ for two-machine operation. Although $I_2=9.3$, the difference between the two values is only 0.1. This may or may not be significant, depending on the relative values of P and nK.

To illustrate the preceding discussion, let us consider a numerical example. A potential customer is considering the use of digital computers for "on-line" application. From field records it is estimated that the mean time to failure of a single machine is 16 hours. The customer wishes to operate his plant for one 8-hour shift per working day. He estimates that the normal operating cost is two hundred dollars per hour for a single machine, and that each hour of down-time will cost him four thousand dollars. Thus, $K=200$, $P=4000$, $\alpha=2$.

To determine n for minimum cost, we find the root of

$$u \ln u = \frac{4000 \times 2}{200} \left\{\ln \left(1/(1 - e^{-1/2})\right\} = 37.32.\right.$$

Then

$$u = 14.1,$$
$$v = P\alpha/Ku = 2.84$$
$$\therefore n = 2.$$

For two machines,

$$\delta_2 = 1 - 2\alpha(1 - e^{-1/\alpha}) + \frac{\alpha}{2}(1 - e^{-2/\alpha})$$
$$= 1 - 1.576 + 0.632 = 0.056.$$

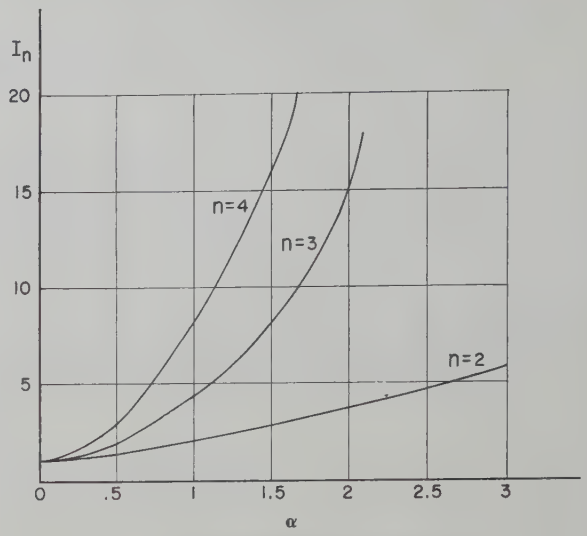


Fig. 1—Improvement factor vs α.

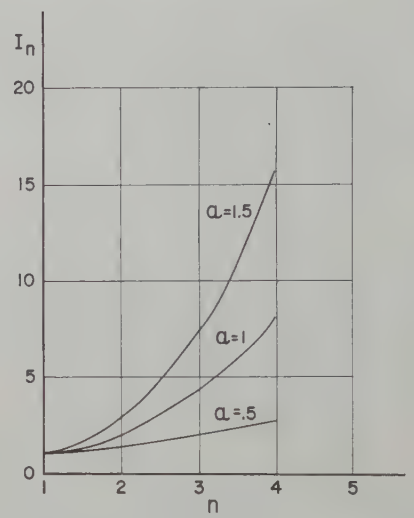


Fig. 2—Improvement factor vs number of machines.

The total cost is:

$$C_2 = 2K + P\delta_2 = 400 + 4000 \times 0.056 = \$624/\text{hr}.$$

The cost improvement by using two machines can be seen by evaluating the cost for single machine operation.

$$\delta_1 = 1 - \alpha(1 - e^{-1/\alpha}) = 0.212$$
$$C_1 = K + P\delta_1 = 200 + 4000 \times 0.212 = \$1048/\text{hr}.$$

There are applications where a fixed amount is paid for any failure during an operating period independent of the down-time associated with the failure. Perhaps the best example of such an application is the case of a digital computer housed in and directing the flight of a missile. Here the operating period is the time of flight of the missile, and any failure during this time causes the entire operation to fail.

The automatic production of precision machine parts where the operating period is completely occupied by

the fabrication of one part is an example of a situation where failure of the on-line computer could ruin the entire output for the period concerned. For this mode of operation, we consider:

Case 3

Our assumptions are:

1) There are n machines operated independently, processing the same input data. The output is taken from a single one of them; if this machine fails, the output is promptly switched to a machine which is operating properly. Once a machine has failed it remains inoperative for the remainder of the working period.

2) All machines run maintenance-free until a time T_0, which corresponds to the end of the working period. All machines are in perfect working order at the beginning of each period.

3) System failure occurs only when all n machines are in the failed condition at the same time.

4) The failure density function for each machine is $f(t) = (1/T_m) \exp(-t/T_m)$, where T_m is the mean time to failure.

5) The cost of each failure is a fixed quantity Q.

For n parallel operated machines the average cost per unit of operating time is:

$$C_n = nK + (Q/T_0)F_n(T_0),$$

where Q/T_0 is the cost per unit of operating time for a system failure and $F_n(T_0)$, usually referred to as the cumulative failure distribution function, is the probability of a system failure at or before the end of the operating period. This equals $[1-R_n(T_0)]$, where $R_n(T_0)$ is the reliability of the n paralleled machines computed over one operating period.

The probability that a single machine will fail at or before a time t is given by

$$F_1(t) = \int_0^t f(t)dt = 1 - e^{-t/T_m}.$$

Therefore

$$R_1(t) = e^{-t/T_m},$$

$$R_1(T_0) = e^{-1/\alpha},$$

and

$$F_1(T_0) = 1 - e^{-1/\alpha}.$$

The probability of a system failure during the operating period is equal to the probability that all n machines fail during this period, which is

$$F_n(T_0) = (1 - e^{-1/\alpha})^n.$$

Therefore

$$R_n(T_0) = 1 - (1 - e^{-1/\alpha})^n.$$

The total cost is

$$C_n = nK + (Q/T_0)F_n(T_0) = nK + (Q/T_0)(1 - e^{-1/\alpha})^n.$$

Using the same procedure as was used in the first and second cases, we find the integer n, which minimizes C_n, to be the greatest integer less than v, where

$$v = 1 + \frac{\frac{1}{\alpha} + \ln \frac{T_0K}{Q}}{\ln (1 - e^{-1/\alpha})}.$$

Reliability as a function of α for $n=1$, 2, 3, and 4 is plotted in Fig. 3. The transition from $n=1$ to $n=2$ produces the greatest increase in reliability, while for $n=4$ the reliability is greater than 0.99 for $T_m/T_0 \geq 3$.

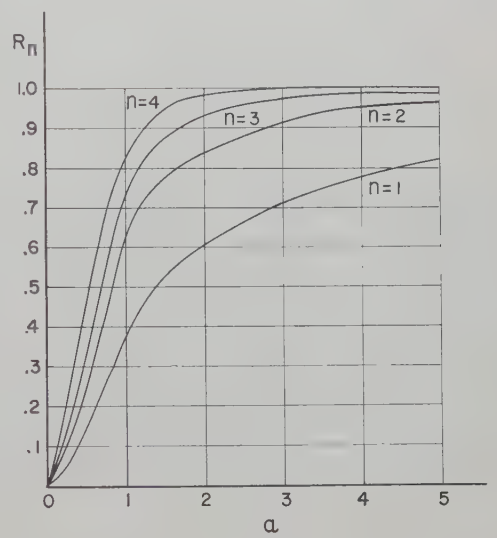


Fig. 3—Reliability vs α.

As a numerical example of this mode of operation, consider the case of an automatic factory which is producing precision machine parts. Assume that the normal operating cost of a single on-line computer is \$100 per hour. Assume further that if the system fails an entire day's production is ruined causing the manufacturer a loss of \$10,000. The operating day is 10 hours and the estimated mean time to failure of the computer is 20 hours. Therefore, $\alpha=2$, and the total cost per hour is:

$$C_n = 10^2n + 10^3(1 - e^{-1/2})^n.$$

The value of n for minimum cost is determined as follows.

$$v = 1 + \frac{\frac{1}{\alpha} + \ln \frac{T_0K}{Q}}{\ln (1 - e^{-1/\alpha})}$$

$$= 1 + \frac{\frac{1}{2} + \ln 0.1}{\ln (1 - e^{-1/2})} = 2.94$$

$$\therefore n = 2$$

For two machines:

$$R_2 = 1 - (1 - e^{-1/2})^2 = 0.845$$

and

$$C_2 = 2 \times 10^2 + 0.155 \times 10^3 = \$355/\text{hr.}$$

For comparison purposes the costs for three machines and for single machine operation are calculated. For three machines:

$$R_3 = 1 - (1 - e^{-1/2})^3 = 0.939$$

and

$$C_3 = 3 \times 10^2 + 0.061 \times 10^3 = \$361/\text{hr.}$$

For one machine:

$$R_1 = 1 - (1 - e^{-1/2}) = 0.606$$

and

$$C_1 = 10^2 + 0.394 \times 10^3 = \$494/\text{hr.}$$

Concluding Remarks

The purpose of this article is to suggest that the cost of a data-processing operation might be reduced by the parallel connection of a number of digital computers. Some simple formulations of the cost per unit operating time as a function of the number of machines have been developed.

We have analyzed three types of cases. For many applications it may not be immediately evident which, if any, of these three cases is appropriate. For example, the difference between Cases 1 and 2 rests on a choice of whether or not to return a failed machine to system operation during the same operating period in which it failed. This decision is, in general a very difficult one; it depends upon the availability of maintenance personnel, the ease of information transfer between machines, the quantity of data to be transferred, the time within the operating period when the failure occurred, the length of time it takes to repair a machine and many other factors.

In some cases, the time remaining within the operating period will be the dominant factor in determining whether or not the system operation should be stopped to permit the return of the repaired machine. (For example, if a machine fails at the very beginning of the day, it may be advisable to attempt to put it back in the system; however, if repairs are completed two minutes before the end of the day, it will not be desirable to stop operation in order to transfer information.) For these cases, therefore, there is a unique time during the operating period before which it will pay to transfer information, and after which a transfer does not pay.

The operating conditions for Cases 2 and 3 are identical; the difference between these cases lies in the nature of the use of the data processing system and the consequent difference in penalty cost evaluation.

Acknowledgment

The authors wish to express their appreciation to R. M. Walker for his contributions and encouragement during the course of this study, and to Betty Flehinger and David Brender for their helpful suggestions.

Boolean Matrix Equations in Digital Circuit Design*

ROBERT S. LEDLEY†

Summary—A systematic digital computational method is given that involves the use of Boolean matrix equations for solving certain types of functional circuit design problems. Specifically, all sets of Boolean functions $f_1(A_1, \cdots, A_I), \cdots, f_J(A_1, \cdots, A_I)$ are found such that if circuits with these outputs are connected to a circuit that generates the known Boolean function $F(f_1, \cdots, f_J, X_1, \cdots, X_K)$, then the output will produce a given desired function $E(A_1, \cdots, A_I, X_1, \cdots, X_K)$. Illustrative examples of the method are presented.

I. Problems Under Consideration

THE kind of digital circuit design problems solved by the methods to be presented in this and subsequent papers can most easily be described by means of a specific example. Fig. 1 is a block diagram of a circuit that has been separated into the circuits f_1, f_2, and F. The inputs to circuit f_1 are the wires labeled A_1, A_2, and A_3, where, of course, the signals produced in f_1 represent some Boolean function of the signals in the inputs, *i.e.*, $f_1 = f_1(A_1, A_2, A_3)$. Similarly for f_2. Now the circuit F has as inputs the wires labeled f_1, f_2, X_1, and X_2, and as output the wire labeled F. The signals produced in F represent some Boolean function of the signals in the inputs, *i.e.*, $F = F(f_1, f_2, X_1, X_2)$. However, since $f_1 = f_1(A_1, A_2, A_3)$ and $f_2(A_1, A_2, A_3)$, we can evidently also describe the signals in F as a Boolean function of A_1, A_2, A_3, X_1, and X_2. For clarity we write this as $E(A_1, A_2, A_3, X_1, X_2)$ where $F(f_1(A_1, A_2, A_3), f_2(A_1, A_2, A_3), X_1, X_2) = E(A_1, A_2, A_3, X_1, X_2)$.

Now that we have described our example, let us see what elementary problems of circuit design can arise here. First, suppose the circuits f_1, f_2, and F have already been constructed. The problem might be to determine the result of wiring them together as in the above diagram, *i.e.*, the problem would be to determine $E(A_1, A_2, A_3, X_1, X_2)$. [See Fig. 2(a).] Second, suppose only the circuits f_1 and f_2 have been constructed, and it is desired to make another circuit F using as inputs f_1, f_2, X_1, and X_2 such that the end result will be the given desired function, $E(A_1, A_2, A_3, X_1, X_2)$. The problem here is to determine the circuit design for F. [See Fig. 2(b).] Third, suppose only the circuit F has been constructed, and it is desired to make circuits f_1 and f_2 so that when wired to F, as in Fig. 1, the final result will be the given desired function $E(A_1, A_2, A_3, X_1, X_2)$. The problem in this case is to design f_1 and f_2. [See Fig. 2(c).] *This third problem is the subject of the present paper.*

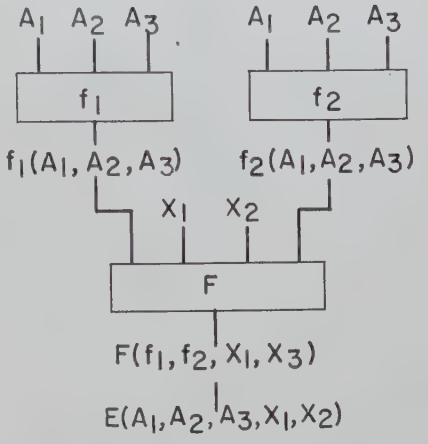


Fig. 1—Example of the situation under consideration in this paper.

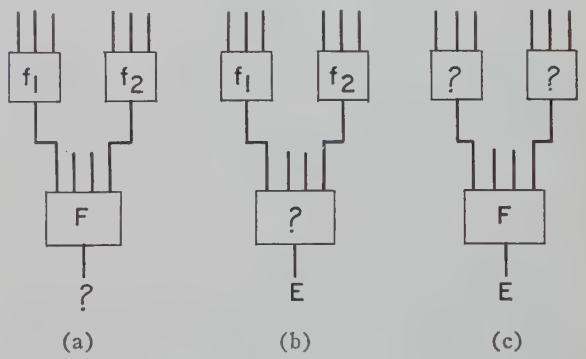


Fig. 2—The three types of circuit design problems that can arise under the circumstances of Fig. 1.

As will be shown, there are many other kinds of digital circuit design problems that can be solved by the methods to be presented. However, this simple picture serves to give a general idea of the kind of problems that can be approached by means of the digital computational methods to be described. For example, in general there can be I input wires $A_1, \cdots, A_I$; K inputs $X_1, \cdots, X_K$; and J functions $f_1, \cdots, f_J$. Hence, in general, $f_s = f_s(A_1, \cdots, A_I)$, $F = F(f_1, \cdots, f_J, X_1, \cdots, X_K)$, and $E = E(A_1, \cdots, A_I, X_1, \cdots, X_K)$.

Brief Notational Review

The notation used in this paper for the usual Boolean algebraic operations of *and*, *or* (inclusive), and *negation*, and their associated gates[1] is summarized in Fig. 3, where A_1 and A_2 are the input wire signals, and f_1 is the output wire signal.

* Manuscript received by the PGEC, December 17, 1958. This work was supported by the Information Systems Branch of the Office of Naval Research and the National Bureau of Standards.

† National Bureau of Standards, Washington, D. C.

[1] It is assumed that the reader is familiar with the conventional method for translating a Boolean algebraic function into a digital circuit diagram and, conversely, a digital circuit diagram into a Boolean algebraic function.

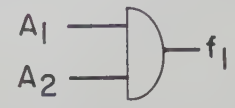
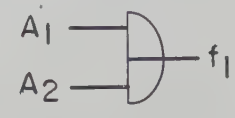
SYMBOLIC NOTATION	CIRCUIT SYMBOL
$f_1 = A_1 \cdot A_2$	
$f_1 = A_1 + A_2$	
$f_1 = \overline{A}_1$	

Fig. 3—Summary of notation for the *and gate, or gate*, and *negation gate* respectively (top to bottom).

Two Boolean functions that are particularly important are $X_r \cdot X_s + \overline{X}_r \cdot \overline{X}_s$ and $\overline{X}_r + X_s$. The former is read, X_r *is equivalent to* X_s (in binary signal value), abbreviated $X_r = X_s$. When $X_r = X_s$ then

$$f(X_r, A_1, A_2, \cdots) = f(X_s, A_1, A_2, \cdots)$$

for any f. The latter (*i.e.*, $\overline{X}_r + X_s$) is read X_r *implies* X_s or *if* X_r *then* X_s, abbreviated $X_r \rightarrow X_s$.

The computational methods for solving the circuit design problems posed above are based on the author's previous digitalization of Boolean algebra in which every Boolean function was associated in a unique way with a binary number, called the designation number of the Boolean function.[2-4] The procedure was to form the designation numbers of the given Boolean functions of a problem; then the computation proceeded with these numbers until the designation number of the solution was obtained. The explicit Boolean function comprising the solution was then derived from its number.

The designation numbers for the *input wire signals*, also called *elementary elements*, are assigned first. Such an assignment is called a *basis*. One such basis for a system of three elementary elements is

$$\begin{aligned} & \textit{0123 4567} \\ \#A_1 &= 0101\ 0101 \\ \#A_2 &= 0011\ 0011 \\ \#A_3 &= 0000\ 1111 \end{aligned}$$

where the upper row of numbers merely signifies the column position. As is clear, the basis is merely a listing of all possible 0 and 1 input signal combinations, of which there are 2^n for n input wires.

The basis given here has the advantage that it may be written down directly, since the first elementary element alternates zeros and units, the second alternates pairs of zeros and units, the third would alternate four zeros and four units, the fourth would alternate eight, etc. The columns of such a basis form binary numbers from 0 to $2^n - 1$, read from left to right, which makes this basis visually convenient. Also note that each column is simply the binary number corresponding to the position of that column. This basis is called the standard basis and, unless otherwise stated, *it is always assumed that all designation numbers refer to this basis.* We shall find occasion to refer to a standard basis as

$$b[A_1, A_2, A_3].$$

Hence this symbol refers to a standard basis where the elementary elements are indicated in their proper order.

The designation number of a Boolean function is merely the horizontal listing of the 0 and 1 signal values corresponding to the respective input possibilities given by the basis. It is clear that the designation number of a Boolean function can be found as follows. To find $\#(A_1 + A_2)$ logically add $\#A_1$ and $\#A_2$ (where logical addition involves no carry and $1+1=1$, $1+0=1$, $0+1=1$, $0+0=0$). For example, with respect to the basis $b[A_1, A_2, A_3]$,

$$\begin{aligned} \#A_1 &= 0101\ 0101 \\ \#A_2 &= 0011\ 0011 \\ \hline \#(A_1 + A_2) &= 0111\ 0111 \end{aligned}.$$

To find $\#(A_1 \cdot A_2)$ logically multiply $\#A_1$ and $\#A_2$ (where logical multiplication involves no carry and $1 \cdot 1 = 1$, $1 \cdot 0 = 0$, $0 \cdot 1 = 0$, $0 \cdot 0 = 0$). For example,

$$\begin{aligned} \#A_1 &= 0101\ 0101 \\ \#A_2 &= 0011\ 0011 \\ \hline \#(A_1 \cdot A_2) &= 0001\ 0001 \end{aligned}.$$

To find $\#(\overline{A}_1)$, we invert each digit of $\#A_1$:

$$\#\overline{A}_1 = 1010\ 1010.$$

Thus, to find the designation number of a Boolean function, we merely perform the indicated operations with respect to the chosen basis. To find $\#(\overline{A}_1 + A_2 \cdot A_3)$:

$$\begin{aligned} \#A_2 &= 0011\ 0011 \\ \#A_3 &= 0000\ 1111 \\ \hline \#(A_2 \cdot A_3) &= 0000\ 0011 \\ \#\overline{A}_1 &= 1010\ 1010 \\ \hline \#(\overline{A}_1 + A_2 \cdot A_3) &= 1010\ 1011 \end{aligned}.$$

Some important results of this digitalization are: 1) $\#X_1 = \#X_2$ if and only if $X_1 = X_2$, and 2) if $X_1 \rightarrow X_2$, then $\#X_2$ has units in at least those positions that correspond to units of $\#X_1$.

[2] R. S. Ledley, "Mathematical foundations and computational methods for a digital logic machine," *J. Operations Res. Soc. Am.*, vol. 2, pp. 249–274; August, 1954.

[3] R. S. Ledley, "Digital computational methods in symbolic logic, with examples in biochemistry," *Proc. Natl. Acad. Sci.*, vol. 41, pp. 498–511; July, 1955.

[4] For a more detailed discussion of the material reviewed in this section see, for example, R. S. Ledley, "Digital Computer and Control Engineering," McGraw-Hill Book Co., Inc., New York, N. Y. (In press.)

Note: There should be no confusion as to the dual role of the symbols $+$, $\cdot$, and $-$. When applied to A_1, A_2, A_3, X_1, X_2, $\cdots$, they are the logical operations "or," "and," and "not"; when applied to designation numbers $\#A_1$, $\#X_1$, $\cdots$, they are the numerical operations "logical sum," "logical multiplication," and "inversion."

Given a designation number, there is a systematic method for finding its symbolic Boolean algebraic functional representation in several different forms. For an example of four different equivalent representations of a given number, consider the designation number 0111 0100: 1) The so-called first canonical form or disjunctive normal form of this number is $A_1 \cdot \overline{A}_2 \cdot \overline{A}_3 + \overline{A}_1 \cdot A_2 \cdot \overline{A}_3 + A_1 \cdot A_2 \cdot \overline{A}_3 + A_1 \cdot \overline{A}_2 \cdot A_3$. 2) The second canonical form or conjunctive normal form is

$$(A_1 + A_2 + A_3) \cdot (A_1 + A_2 + \overline{A}_3) \cdot (A_1 + \overline{A}_2 + \overline{A}_3) \cdot (\overline{A}_1 + \overline{A}_2 + \overline{A}_3).$$

3) The simplest sum of products form (*i.e.*, a sum of products that has the least number of operations $+$ and $\cdot$ is $A_1 \cdot \overline{A}_2 + A_2 \cdot \overline{A}_3$. 4) The simplest product of sums form is $(A_1 + A_2) \cdot (\overline{A}_2 + \overline{A}_3)$. (See for instance Ledley,[2] McCluskey,[5] and Harris,[6] for systematic methods of developing these functional representations from designation numbers.)

In the advanced computational methods which are the subject of the present papers, the computations are not always carried out in terms of the designation numbers. Rather, the designation numbers are turned into Boolean matrices; the computations involve these matrices; the resulting matrix is then changed back into a designation number and the explicit Boolean function comprising the solution is derived from its number. This process is summarized in Fig. 4.

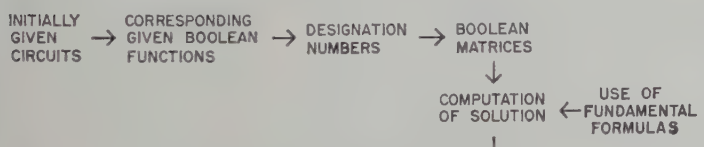


Fig. 4—Stages in the solution of a problem by the methods of this paper.

II. Antecedence and Consequence Solutions to Boolean Equations

General Formulation

The method of solving digital circuit design problems under consideration in this paper is closely related to the notion of antecedence and consequence solutions for Boolean equations.[3] We shall first formulate the meaning of antecedence and consequence solutions in terms of logic, and follow this discussion by the interpretation of such solutions in terms of digital computer circuits.

The majority of logical problems involve a given or accepted premise, hypothesis, rule, or other logical relationship which is the *given equation* and essentially comprises the statement of the problem. There are two types of solutions to a set of given equations, the *antecedence solutions* and the *consequence solutions*. Antecedence solutions are hypotheses or theories from which the given equation can be deduced; consequence solutions can be deduced from the given equation. In other words, the truth of the antecedence solutions is sufficient for the truth of the given equation, but the truth of the consequence solutions is necessary for the truth of the given equation. If the given equation is true, then the consequence solutions are true, while the antecedence solutions may or may not be true; but the truth of the given equation can be deduced from the hypotheses embodied in the antecedence solutions, this latter being the method for theory construction.

However, to produce just one antecedence or one consequence solution to a given equation is usually trivial; hence logical problems usually require solutions of a *specified form* or solutions involving only *certain specified elementary elements*, or both. Occasionally solutions with the required properties do not exist, and the problem is extended to determine under what conditions such solutions do exist for the given equation.

Interpretation in Terms of Circuit Design

In terms of the design of digital circuitry, the given equation usually represents the Boolean function corresponding to the final desired output of the circuit being designed. That is to say, the goal of the circuit design process is to produce a circuit the output of which is represented by the Boolean function embodied in the given equation. The given equation therefore tells what combinations of input conditions are to result in a unit output voltage signal. A consequence solution of the given equation is any Boolean function representing a circuit design that will produce a unit output signal for *at least all* input conditions for which the given equation is to produce such a signal; in fact, it may produce a unit output signal for other input conditions as well. An antecedence solution of the given equation is any Boolean function representing a circuit design that will produce a *zero* output signal for *at least all* input conditions for which the given equation is to produce such a signal; in fact, it may produce a zero output signal for other input conditions as well. In other words, antecedence and consequence solutions to the given equation do not necessarily represent circuits that are equivalent to that of the given equation, but do represent circuits that are "close" to that of the given equation. However, a single solution that is *both* an antecedence *and* consequence solution does represent a circuit that is equivalent to that of the given equation.

[5] E. J. McCluskey, Jr., "Minimization of Boolean functions," *Bell Sys. Tech. J.*, vol. 35, pp. 1417–1444; November, 1956.

[6] B. Harris, "An algorithm for determining minimal representations of a logic function," IRE Trans. on Electronic Computers, vol. EC-6, pp. 103–108; June, 1957.

Let us now describe the problem in more specific mathematical terms and give particular instances of its application to the design of digital circuitry. Suppose that the given equation is a Boolean function of the elementary elements $A_1, A_2, \cdots, A_I, X_1, X_2, \cdots, X_K$, and hence can represented as

$$E(A_1, A_2, \cdots, A_I, X_1, X_2, \cdots, X_K).$$

Suppose, also, an antecedence solution is desired that is a specified function of $f_1, f_2, \cdots, f_J, X_1, X_2, \cdots, X_K$, where each f_J is itself a function of $A_1, A_2, \cdots, A_I$. We can then represent an antecedence solution as

$$F_a(f_1, f_2, \cdots, f_J, X_1, X_2, \cdots, X_K)$$

where $f_1=f_1(A_1, \cdots, A_I)$, $f_2=f_2(A_1, \cdots, A_I), \cdots, f_J=f_J(A_1, \cdots, A_I)$. In order that F_a be an antecedence solution, we must have

$$F_a \rightarrow E$$

(where it may happen that $F_a=E$). Consequence solutions are similarly defined to be of the form $F_c(f_1, f_2, \cdots, f_J, X_1, X_2, \cdots, X_K)$ where $f_s=f_s(A_1, \cdots, A_I)$. But in order that it be a consequence solution, we must have

$$E \rightarrow F_c$$

(where it may happen that $E=F_c$). If F_c is the same as F_a, *i.e.*, if for the same F both $F \rightarrow E$ and $E \rightarrow F$, then $E=F$. These are the ingredients that enter into the problem as summarized in Fig. 5.

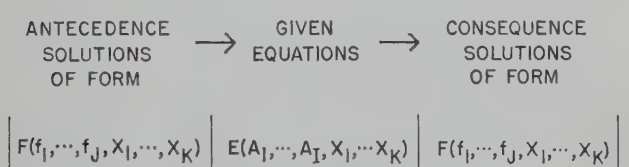


Fig. 5—Relation between antecedence and consequence solutions.

III. The Computational Method

Three Steps

In the problems we are considering the functions

$$E(A_1, \cdots, A_I, X_1, \cdots, X_K)$$

and

$$F(f_1, \cdots, f_J, X_1, \cdots, X_K)$$

are *given* explicitly in the statement of the problem; it is desired to find the unknown functions

$$f_1(A_1, \cdots, A_I), \cdots, f_J(A_1, \cdots, A_I).$$

The first step in the computational method is to systematically derive Boolean matrices (E_{ki}) and (F_{jk}) corresponding to the given functions E and F respectively. The second step is to use the Boolean matrix equation formulas that will be given below to determine the matrix (R_{ji}). The third step is to obtain the desired set of functions $f_1, \cdots, f_J$ from the computed matrix (R_{ji}).

The first step: The Boolean matrices under consideration in this paper are rectangular arrays with elements that can be only 0 or 1. The Boolean matrices corresponding to E and F can easily be determined directly from $\#E$ and $\#F$ respectively as follows: partition $\#E$ (or $\#F$) into 2^K parts each of which contains 2^I (or 2^J) bits; then the *rows* of (E_{ki}) are the successive parts of $\#E$, while the *columns* of (F_{jk}) are the successive parts of $\#F$. For example suppose the function E is given by

$$\begin{aligned} E\colon\ & (\overline{A}_1 \cdot \overline{A}_3 + \overline{A}_2 \cdot A_3) \cdot \overline{X}_1 \cdot \overline{X}_2 \\ & + (\overline{A}_1 \cdot \overline{A}_3 + A_1 A \cdot_2) \cdot X_1 \cdot \overline{X}_2 \\ & + (A_1 \cdot \overline{A}_3 + A_2 \cdot A_3) \cdot \overline{X}_1 \cdot X_2 \\ & + (A_1 \cdot \overline{A}_2 \cdot \overline{A}_3 + \overline{A}_1 \cdot A_2 \cdot A_3) \cdot X_1 \cdot X_2. \end{aligned}$$

With respect to $b[A_1, A_2, A_3, X_1, X_2]$, find

$$\begin{array}{lcccc} \#E = & 1010\ 1100 & 1011\ 0001 & 0101\ 0011 & 0100\ 0010 \\ i & 0123\ 4567 & 0123\ 4567 & 0123\ 4567 & 0123\ 4567 \\ k & 0 & 1 & 2 & 3 \end{array}$$

whence

$$(E_{ki}) = \begin{pmatrix} 1010\ 1100 \\ 1011\ 0001 \\ 0101\ 0011 \\ 0100\ 0010 \end{pmatrix}.$$

Note that the (E_{ki}) matrix turns out to be essentially the Veitch Chart[7] for the function E, where the 2^I combinations of input signals of the $A_1, \cdots, A_I$ label the columns, and the 2^K combinations of input signals of the $X_1, \cdots, X_K$ label the rows. For our function E given above we have

$$\begin{array}{lc} & 0123\ 4567\ i \\ A_1 & 0101\ 0101 \\ A_2 & 0011\ 0011 \\ A_3 & 0000\ 1111 \end{array}$$

$$(E_{ki}) = \begin{array}{|c|cc|c} 1010\ 1100 & 0 & 0 & 0 \\ 1011\ 0001 & 1 & 0 & 1 \\ 0101\ 0011 & 0 & 1 & 2 \\ 0100\ 0010 & 1 & 1 & 3 \\ \hline & X_1 & X_2 & k \end{array}.$$

As another example suppose that the function F is given by

$$F\colon \bar{f}_1 \cdot \bar{f}_2 \cdot \overline{X}_2 + f_1 \cdot \bar{f}_2 \cdot X_2 + \bar{f}_1 \cdot f_2 \cdot (X_1 \cdot \overline{X}_2 + \overline{X}_1 \cdot X_2) + f_1 \cdot f_2 \cdot \overline{X}_1 \cdot \overline{X}_2.$$

[7] E. W. Veitch, "A chart method for simplifying truth functions," *Proc. Assoc. Comp. Mach.* (Pittsburgh, Pa. meeting), pp. 127–133; May 2–3, 1952.

With respect to $b[f_1, f_2, X_1, X_2]$ we find that

$$\begin{array}{rcccc} \#F = & 1001 & 1010 & 0110 & 0100 \\ j & 0123 & 0123 & 0123 & 0123 \\ k & 0 & 1 & 2 & 3 \end{array}$$

whence

$$(F_{jk}) = \begin{pmatrix} 1100 \\ 0011 \\ 0110 \\ 1000 \end{pmatrix}.$$

The matrix (F_{jk}) can likewise be considered to be a Veitch Chart, where the 2^J combinations of input signals of the $f_1, \cdots, f_J$ label the rows, and the 2^K combinations of input signals of the $X_1, \cdots, X_K$ label the columns.

For our function F given above we have

$$\begin{array}{rl} & 0123\ k \\ X_1 & 0101 \\ X_2 & 0011 \end{array}$$

$$(F_{jk}) = \begin{array}{l} \\ \begin{pmatrix} 1100 \\ 0011 \\ 0110 \\ 1000 \end{pmatrix} \end{array} \begin{array}{c} j \\ \begin{matrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{matrix} \\ f_1\ f_2 \end{array}.$$

The second step: When the matrices (E_{ki}) and (F_{jk}) have been derived from the functions E and F given by the statement of the problem, they are substituted into the following two matrix formulas[8] to compute the two matrices $(R_{ji})_a$ and $(R_{ji})_c$:

Antecedence	Consequence
$(F_{jk}) \otimes (\overline{E}_{ki}) = (\overline{R}_{ji})_a$	$(\overline{F}_{jk}) \otimes (E_{ki}) = (\overline{R}_{ji})_c$

where the Boolean matrix multiplication operation indicated by $\otimes$ represents ordinary matrix multiplication, except that logical addition and logical multiplication of the elements now replaces ordinary addition and multiplication, *i.e.*, if $(M_{jk}) \otimes (N_{ki}) = (L_{ji})$ then an element

$$L_{ji} = \sum_{k=0}^{K-1} M_{jk} \cdot N_{ki}$$

where $\sum$ represents logical addition.

Also if (M_{pq}) is a Boolean matrix of this type, then by $(\overline{M}_{pq})$ we mean the matrix formed by changing all zero elements of (M_{pq}) to units, and all unit elements of (M_{pq}) to zeros.

For example, using our functions E and F from above we find

$$(F_{jk}) \otimes (\overline{E}_{ki}) = \begin{pmatrix} 1100 \\ 0011 \\ 0110 \\ 1000 \end{pmatrix} \otimes \begin{pmatrix} 0101 & 0011 \\ 0100 & 1110 \\ 1010 & 1100 \\ 1011 & 1101 \end{pmatrix} = \begin{pmatrix} 0101 & 1111 \\ 1011 & 1101 \\ 1110 & 1110 \\ 0101 & 0011 \end{pmatrix} = (\overline{R}_{ji})_a.$$

Therefore

$$(R_{ji})_a = \begin{pmatrix} 1010 & 0000 \\ 0100 & 0010 \\ 0001 & 0001 \\ 1010 & 1100 \end{pmatrix}.$$

$$(\overline{F}_{jk}) \otimes (E_{ki}) = \begin{pmatrix} 0011 \\ 1100 \\ 1001 \\ 0111 \end{pmatrix} \otimes \begin{pmatrix} 1010 & 1100 \\ 1011 & 0001 \\ 0101 & 0011 \\ 0100 & 0010 \end{pmatrix} = \begin{pmatrix} 0101 & 0011 \\ 1011 & 1101 \\ 1110 & 1110 \\ 1111 & 0011 \end{pmatrix} = (\overline{R}_{ji})_c.$$

Thus

$$(R_{ji})_c = \begin{pmatrix} 1010 & 1100 \\ 0100 & 0010 \\ 0001 & 0001 \\ 0000 & 1100 \end{pmatrix}.$$

The third step: As we will now show, sets of functions can be derived from the matrices $(R_{ji})_a$ and $(R_{ji})_c$. The sets of functions derived from $(R_{ji})_a$ when substituted into F, produce antecedence solutions; the sets derived from $(R_{ji})_c$ produce consequence solutions. Sets of functions that turn out to be *both antecedence and consequence solutions*, when substituted into F will make $F=E$, *and are therefore the desired solutions for our digital circuit design problem.* The process of generating $f_1(A_1, \cdots, A_I), f_2(A_1, \cdots, A_I), \cdots, f_J(A_1, \cdots, A_I)$ from a given (R_{ji}) is not unique, in contrast to the other processes described in this section. Several sets of f_s are often possible while, on the other hand, it can happen that no sets f_s correspond to a given (R_{ji}). Two bases are consulted in this process, namely $b[f_1, \cdots, f_J]$ and $b[A_1, \cdots, A_I]$, both written in the usual pattern. The desired solution is computed by means of the result array which consists of an as yet empty array with 2^I columns and J rows; the columns are indexed in order by i from left to right, the rows corresponding to $f_1, \cdots, f_J$, assigned from top to bottom. Considering those pairs of indexes j, i for which $R_{ji}=1$, we place the jth column of $b[f_1, \cdots, f_J]$ in the ith column of the result array. The rows of the result array thus filled are the designation numbers of the corresponding $f_1, \cdots, f_J$ with respect to the basis $b[A_1, \cdots, A_I]$.

For example consider

$$(R_{ji})_a = \begin{pmatrix} 1010 & 0000 \\ 0100 & 0010 \\ 0001 & 0001 \\ 1010 & 1100 \end{pmatrix}$$

[8] The proof of these formulas will be developed in a later paper.

$b[f_1, f_2]$:

j 0123

$\#f_1 = 0101$

$\#f_2 = 0011$

Result Array:

i	0	1	2	3	4	5	6	7
$\#f_1(A_1, A_2, A_3)$	0, 1	1	0, 1	0	1	1	1	0
$\#f_2(A_1, A_2, A_3)$	0, 1	0	0, 1	1	1	1	0	1

There are precisely four possible sets of solutions, which with respect to $b[A_1, A_2, A_3]$ are seen to be

$$\#f_1 = 0100\ 1110 = \#(A_1 \cdot \overline{A}_2 + \overline{A}_1 \cdot A_3),$$
$$\#f_2 = 0001\ 1101 = \#(A_1 \cdot A_2 + \overline{A}_2 \cdot A_3),$$

$$\#f_1 = 0110\ 1110 = \#A_1 \cdot \overline{A}_2 + \overline{A}_1 \cdot A_2 + \overline{A}_1 \cdot A_3),$$
$$\#f_2 = 0011\ 1101 = \#(A_2 \cdot \overline{A}_3 + \overline{A}_2 \cdot A_3 + A_1 \cdot A_2),$$

$$\#f_1 = 1100\ 1110 = \#(\overline{A}_2 + \overline{A}_1 \cdot A_3),$$
$$\#f_2 = 1001\ 1101 = \#(A_1 \cdot A_2 + \overline{A}_1 \cdot \overline{A}_2 + \overline{A}_2 \cdot A_3),$$

$$\#f_1 = 1110\ 1110 = \#(\overline{A}_1 + \overline{A}_2),$$
$$\#f_2 = 1011\ 1101 = \#(\overline{A}_1 \cdot \overline{A}_3 + A_1 \cdot A_2 + \overline{A}_2 \cdot A_3).$$

However, to find all solutions that are *both* antecedence and consequence solutions, we do not need to generate all sets f_s corresponding to $(R_{ji})_a$ and all sets f_s corresponding to $(R_{ji})_c$ and compare. Rather to find those solutions that are both antecedence and consequence solutions, form the matrix (R_{ji}) composed of all units *common to both* $(R_{ji})_a$ and $(R_{ji})_c$, *i.e.*,

$$(R_{ji}) = (R_{ji})_a \cdot (R_{ji})_c$$

where the symbol $\cdot$ represents logical multiplication of the corresponding elements of the matrices, *i.e.*, if $(L_{ji}) = (M_{ji}) \cdot (N_{ji})$ then $L_{ji} = M_{ji} \cdot N_{ji}$. For example, for $(R_{ji})_a$ and $(R_{ji})_c$ computed above we have

$$(R_{ji}) = (R_{ji})_a \cdot (R_{ji})_c = \begin{pmatrix} 1010 & 0000 \\ 0100 & 0010 \\ 0001 & 0001 \\ 0000 & 1100 \end{pmatrix}$$

whence there is only one set of solutions that are both antecedence and consequence solutions, namely

$$f_1 = A_1 \cdot \overline{A}_2 + \overline{A}_1 \cdot A_3$$
$$f_2 = A_1 \cdot A_2 + \overline{A}_2 \cdot A_3.$$

IV. Examples

In this section three examples will be given. The first illustrates a normally complicated situation; the second illustrates a situation where many solutions turn up and additional criteria are required; and the third illustrates the design of a multiple output circuit. An analogy can be made here with the learning of the calculus: It is not enough in itself; one must also learn how to apply the calculus to various kinds of problems. Of course, as in the calculus, practical application of the methods in engineering research ultimately depends on the ingenuity and skill of the design engineer.

First Example

Suppose a circuit F has already been constructed with inputs f_1, f_2, X_1, and X_2 such that its output Boolean function is given by

$$F: \bar{f}_1 \cdot \bar{f}_2 \cdot \overline{X}_2 + f_1 \cdot \bar{f}_2 \cdot X_2 + \bar{f}_1 \cdot f_2 \cdot (X_1 \cdot \overline{X}_2 + \overline{X}_1 \cdot X_2) + f_1 \cdot f_2 \cdot \overline{X}_1 \cdot \overline{X}_2.$$

Suppose it was desired to produce—by attaching additional circuits to the f_1 and f_2 inputs of this F circuit—the following output result:

$$E = (\overline{A}_1 \cdot \overline{A}_3 + \overline{A}_2 \cdot A_3) \cdot \overline{X}_1 \cdot \overline{X}_2 + (\overline{A}_1 \cdot \overline{A}_3 + A_1 \cdot A_2) \cdot X_1 \cdot \overline{X}_2 + (A_1 \cdot \overline{A}_3 + A_2 \cdot A_3) \cdot \overline{X}_1 \cdot X_2 + (A_1 \cdot \overline{A}_2 \cdot \overline{A}_3 + \overline{A}_1 \cdot A_2 \cdot A_3) \cdot X_1 \cdot X_2.$$

The problem is to design the two circuits with inputs A_1, A_2, and A_3 whose outputs should be connected to f_1 and f_2, respectively.

Solution: The first step is to derive the matrices (E_{ki}) and (F_{jk}) from the functions E and F. For the functions E and F of the first example, we have already found the corresponding matrices in Section III. The second step is to determine $(R_{ji})_a$ and $(R_{ji})_c$ by substituting (E_{ki}) and (F_{jk}) into the antecedence and consequence formulas. For this example, we have determined these matrices in Section III. The third step is to determine the explicit solution to our circuit design problem from $(R_{ji})_a$ and $(R_{ji})_c$. This is accomplished by calculating (R_{ji}), the matrix containing only the units common to $(R_{ji})_a$ and $(R_{ji})_c$, and determining the sets of functions f_s corresponding to (R_{ji}). We have completed this calculation for this example in Section III, and found that there results a single set of solutions, namely

$$f_1 = A_1 \cdot \overline{A}_2 + \overline{A}_1 \cdot A_3$$
$$f_2 = A_1 \cdot A_2 + \overline{A}_2 \cdot A_3.$$

The desired circuits are therefore as in Fig. 6.

Second Example

It sometimes happens that there are many antecedence or consequence solutions. This merely means that additional criteria are needed to choose the desired solution. Such criteria are that the desired solutions be both antecedence and consequence solutions (so that $E = F$), that the set of functions of the solution be as simple as possible, that the set of functions be independent, that the set of functions satisfy some given constraints, etc. These additional criteria will quickly reduce the number of solutions. This is actually an exceedingly interesting

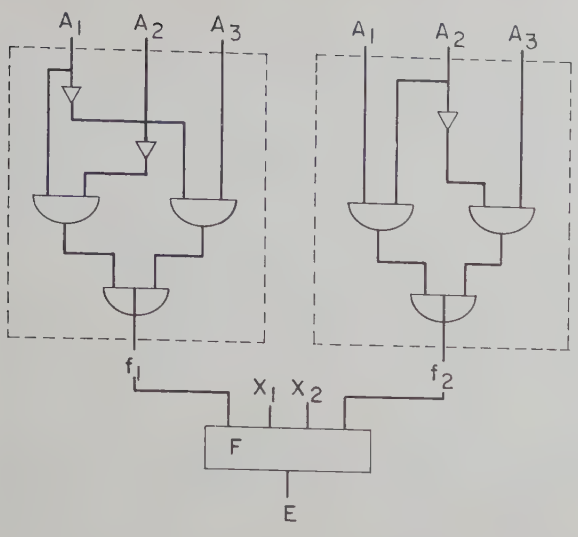


Fig. 6—Circuit design solution for the *First Example.*

point, for here the mathematics is essentially disclosing whether or not the original formulation of the problem was adequately precise.

To illustrate these points, consider the following problem. Suppose it is desired to construct a circuit with output function E:

$$A_1 \cdot A_2 + X_1 \cdot \overline{A}_2 \cdot A_3 + X_2 \cdot (A_1 \cdot \overline{A}_3 + A_2).$$

Suppose, in addition, a certain circuit had already been constructed, namely one with input wires f_1, f_2, f_3, X_1, and X_2, the output of which is F:

$$\bar{f}_1 \cdot f_3 + X_1 \cdot f_1 \cdot \bar{f}_2 + X_2 \cdot f_2 \cdot \bar{f}_3.$$

It is desired to construct three digital circuits to be connected directly to the inputs f_1, f_2 and f_3 of the already completed circuit F. Each of these three circuits has inputs A_1, A_2, and A_3, and the problem is to design the three circuits so that when connected to the inputs of F the total resultant output will be the function E.

Solution: With respect to $b[A_1, A_2, A_3, X_1, X_2]$ and $b[f_1, f_2, f_3, X_1, X_2]$, respectively, we find

$$\#E = 0001\ 0001\ 0001\ 1101\ 0111\ 0011\ 0111\ 1111$$

and

$$\#F = 0000\ 1010\ 0100\ 1110\ 0011\ 1010\ 0111\ 1110$$

whence

$$(E_{ki}) = \begin{pmatrix} 0001 & 0001 \\ 0001 & 1101 \\ 0111 & 0011 \\ 0111 & 1111 \end{pmatrix} \quad \text{and} \quad (F_{jk}) = \begin{pmatrix} 0000 \\ 0101 \\ 0011 \\ 0011 \\ \\ 1111 \\ 0101 \\ 1111 \\ 0000 \end{pmatrix}.$$

Substituting in the antecedence formula we have

$$(F_{jk}) \otimes (\overline{E}_{ki})$$

$$= \begin{pmatrix} 0000 \\ 0101 \\ 0011 \\ 0011 \\ \\ 1111 \\ 0101 \\ 1111 \\ 0000 \end{pmatrix} \otimes \begin{pmatrix} 1110 & 1110 \\ 1110 & 0010 \\ 1000 & 1100 \\ 1000 & 0000 \end{pmatrix} = \begin{pmatrix} 0000 & 0000 \\ 1110 & 0010 \\ 1000 & 1100 \\ 1000 & 1100 \\ \\ 1110 & 1110 \\ 1110 & 0010 \\ 1110 & 1110 \\ 0000 & 0000 \end{pmatrix} = (\overline{R}_{ji})_a.$$

Thus since in $(R_{ji})_a$ there are 2 units in the zeroth column, 4 units in the first column, 4 units in the second column, 8 units in the third column, etc., then there are precisely $2 \cdot 4 \cdot 4 \cdot 8 \cdot 4 \cdot 4 \cdot 4 \cdot 8 = 131{,}072$ non-equivalent sets of antecedence solutions. This seems to be a large number of possible solutions, but, after all, it was narrowed down from the $8^8 = 16{,}777{,}216$ total possible sets of three functions of three variables! However, we are interested in solutions such the $F = E$, and our antecedence solutions only assure us that $F \rightarrow E$. Therefore, let us find the consequence solutions such that $E \rightarrow F$, and then any solutions that turn out to be *both* antecedence and consequence solutions will satisfy $F = E$, *i.e.*, both $F \rightarrow E$ and $E \rightarrow F$.

For consequence solutions we have

$$(\overline{F}_{jk}) \otimes (E_{ki})$$

$$= \begin{pmatrix} 1111 \\ 1010 \\ 1100 \\ 1100 \\ \\ 0000 \\ 1010 \\ 0000 \\ 1111 \end{pmatrix} \otimes \begin{pmatrix} 0001 & 0001 \\ 0001 & 1101 \\ 0111 & 0011 \\ 0111 & 1111 \end{pmatrix} = \begin{pmatrix} 0111 & 1111 \\ 0111 & 0011 \\ 0001 & 1101 \\ 0001 & 1101 \\ \\ 0000 & 0000 \\ 0111 & 0011 \\ 0000 & 0000 \\ 0111 & 1111 \end{pmatrix} = (\overline{R}_{ji})_c$$

whence

$$(R_{ji}) = (R_{ji})_a \cdot (R_{ji})_c = \begin{pmatrix} 1000 & 0000 \\ 0000 & 1100 \\ 0110 & 0010 \\ 0110 & 0010 \\ \\ 0001 & 0001 \\ 0000 & 1100 \\ 0001 & 0001 \\ 1000 & 0000 \end{pmatrix}.$$

The result array becomes

i	0	1	2	3	4	5	6	7
$\#f_1(A_1, A_2, A_3)$	0, 1	0, 1	0, 1	0, 0	1, 1	1, 1	0, 1	0, 0
$\#f_2(A_1, A_2, A_3)$	0, 1	1, 1	1, 1	0, 1	0, 0	0, 0	1, 1	0, 1
$\#f_3(A_1, A_2, A_3)$	0, 1	0, 0	0, 0	1, 1	0, 1	0, 1	0, 0	1, 1

Thus the number of desired solutions has been reduced from 131,072 to $2\cdot2\cdot2\cdot2\cdot2\cdot2\cdot2\cdot2=256$ solutions. But this is still too many. Hence we can add another criterion to our desired solution, *e.g.*, that it be a "simple" solution.

To choose a simple solution from the 256 possible solutions, let us see if any set of solutions has an elementary element as one of the functions. *Note* that f_1 can be chosen to be an elementary element, namely $\overline{A}_2$. Throwing away the other solutions we have:

Result array:

i	0	1	2	3	4	5	6	7	
$\#f_1(A_1, A_2, A_3)$	1	1	0	0	1	1	0	0	$=\#\overline{A}_2$
$\#f_2(A_1, A_2, A_3)$	1	1	1	0, 1	0	0	1	0, 1	
$\#f_3(A_1, A_2, A_3)$	1	0	0	1	0, 1	0, 1	0	1	

But we still have a choice of $2\cdot2\cdot2\cdot2=16$ sets of solutions with $f_1=\overline{A}_2$ because of columns 3, 4, 5, and 7. Let us try to choose f_2 as easily as possible. *Note* that we can choose $\#f_2(A_1, A_2, A_3)=1111\ 0011=A_2+\overline{A}_3$, and then $\#f_3=1001\ 1001=A_1\cdot A_2+\overline{A}_1\cdot\overline{A}_2$. Hence, a simple set of solutions is:

Result array:

i	0	1	2	3	4	5	6	7	
$\#f_1(A_1, A_2, A_3)$	1	1	0	0	1	1	0	0	$=\#\overline{A}_2$
$\#f_2(A_1, A_2, A_3)$	1	1	1	1	0	0	1	1	$=\#A_2+\overline{A}_3$
$\#f_3(A_1, A_2, A_3)$	1	0	0	1	1	0	0	1	$=\#A_1\cdot A_2+\overline{A}_1\cdot\overline{A}_2$.

Hence the three desired circuits are as shown in Fig. 7 (opposite).[9]

Third Example

Suppose it is desired to construct a circuit with the following three multiple outputs:

$$E_1 = A_1\cdot(\overline{A}_2 + A_3\cdot\overline{A}_4) + \overline{A}_2\cdot A_3$$

$$E_2 = (\overline{A}_1\cdot\overline{A}_3 + A_1\cdot A_3)\cdot(\overline{A}_2 + \overline{A}_4)$$

$$E_3 = \overline{A}_2\cdot(\overline{A}_3 + A_1\cdot A_4) + (A_1 + A_2)\cdot A_3\cdot\overline{A}_4.$$

Suppose in addition that it is desired to utilize the following three already constructed circuits:

$$F_1 = (f_1 + \bar{f}_2)\cdot\bar{f}_3 + f_1\cdot\bar{f}_2$$

$$F_2 = f_1\cdot\bar{f}_3 + \bar{f}_1\cdot f_3$$

$$F_3 = f_2\cdot\bar{f}_3 + f_2\cdot f_3.$$

The problem is to design the three circuits

$$f_1 = f_1(A_1, A_2, A_3, A_4), \quad f_2 = f_2(A_1, A_2, A_3, A_4)$$

and

$$f_3 = f_3(A_1, A_2, A_3, A_4)$$

such that when these are connected to F_1, F_2, and F_3, the outputs of these circuits will be respectively E_1, E_2, and E_3. (See Fig. 8.)

Solution: Solve for $(R_{ji})_1$, $(R_{ji})_2$, and $(R_{ji})_3$, respectively, considering first E_1 and F_1, next E_2 and F_2, and finally E_3 and F_3. Then the desired (R_{ji}) is the matrix composed of all units common to these three matrices, *i.e.*,

$$(R_{ji}) = (R_{ji})_1\cdot(R_{ji})_2\cdot(R_{ji})_3.$$

[9] Other approaches also yield simple solutions; for instance another simple solution to the *second example* is $f_1=\bar{A}_2\cdot A_3$, $f_2=A_2+A_1\cdot\bar{A}_3$, $f_3=A_1\cdot A_2$.

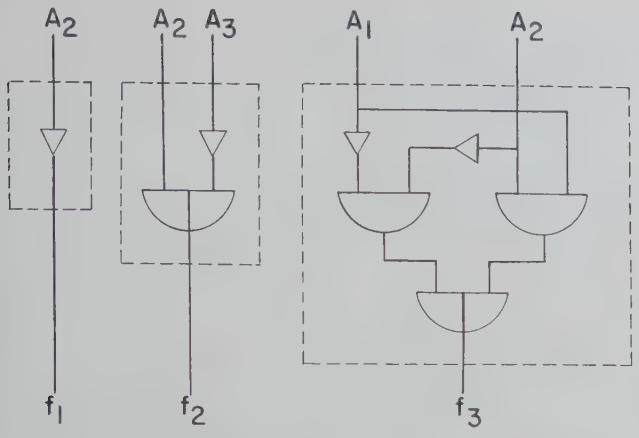


Fig. 7—Circuit design solution for the *Second Example*.

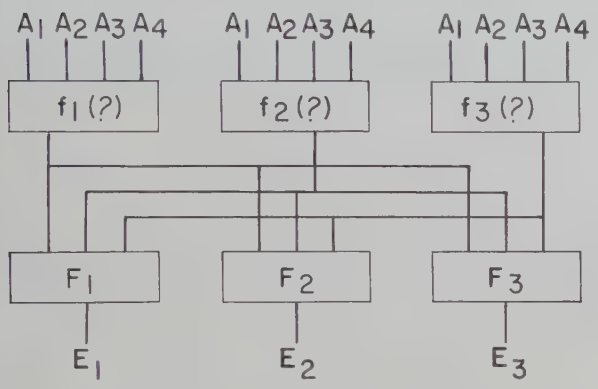


Fig. 8—The problem of the *Third Example*.

Carrying through the details, it is found that

$$(R_{ji}) = \begin{pmatrix} 0000 & 1000 & 0000 & 1000 \\ 0000 & 0000 & 0000 & 0000 \\ 0000 & 0010 & 0000 & 0000 \\ 0000 & 0101 & 0000 & 0100 \\ & & & \\ 1000 & 0000 & 1000 & 0000 \\ 0100 & 0000 & 0100 & 0000 \\ 0010 & 0000 & 0000 & 0000 \\ 0001 & 0000 & 0011 & 0011 \end{pmatrix}$$

whence

$$\#f_1 = 0101\ 0101\ 0111\ 0111 = \#(A_1 + A_2 \cdot A_4$$

$$\#f_2 = 0011\ 0111\ 0011\ 0111 = \#(A_2 + A_3 \cdot A_1)$$

$$\#f_3 = 1111\ 0000\ 1111\ 0011 = \#(\overline{A}_3 + A_4 \cdot A_2).$$

Acknowledgment

The author wishes to express his appreciation for the encouragement and advice of R. D. Elbourn of the National Bureau of Standards.

The Residue Number System*

HARVEY L. GARNER†

Summary—A novel number system called the residue number system is developed from the linear congruence viewpoint. The residue number system is of particular interest because the arithmetic operations of addition, subtraction and multiplication may be executed in the same period of time without the need for carry. The main difficulties of the residue code pertain to the determination of the relative magnitude of two residue representations, and to the division process. A discussion of the arithmetic operations and the conversion process required to convert from a residue code to a weighted code is given. It is concluded that in its present state the residue code is probably not suitable for general purpose computation but is suitable for a special class of control problems. Further research in both components and arithmetic is required if a residue code suitable for general purpose computation is to be obtained.

Introduction

IN THIS PAPER we develop and investigate the properties of a novel system called the residue code or residue number system. The residue number system is of particular interest because the arithmetic operations of addition and multiplication may be executed in the same time as required for an addition operation. The main difficulty of the residue code relative to arithmetic operations is the determination of the relative magnitude of two numbers expressed in the residue code. The residue code is probably of little utility for general purpose computation, but the code has many characteristics which recommend its use for special purpose computations. The residue code is most easily developed in terms of linear congruences. A brief discussion of the pertinent properties of congruences is presented in the next section.

Congruences

The congruence relationship is expressed as

$$A \equiv \alpha \text{ Mod } b \tag{1}$$

which is read, A is congruent to α Modulo b. The congruence states that the equation

$$A = \alpha + bt \tag{2}$$

is valid for some value of t, where A, α, b and t are integers. α is called the residue and b the base or modulus of the number A.

As examples of congruences consider

$$\begin{aligned} 10 &\equiv 7 \text{ Mod } 3 \\ 10 &\equiv 4 \text{ Mod } 3 \\ 10 &\equiv 1 \text{ Mod } 3. \end{aligned} \tag{3}$$

In these examples the integers 7, 4, and 1 form a residue class of 10 Mod 3. Of particular importance is the least positive residue of the class which in this example is one. The least positive residue is that residue for which $0 \leq \alpha < b$.

Consider the following set of congruences. Given

$$\begin{aligned} A_1 &\equiv \alpha_1 \text{ Mod } b \\ &\vdots \\ A_n &\equiv \alpha_n \text{ Mod } b. \end{aligned} \tag{4}$$

Then:

1) Congruences with respect to the same modulus may be added and the result is a valid congruence.

$$\sum_{i=1}^{n} A_i \equiv \left(\sum_{i=1}^{n} \alpha_i \right) \text{Mod } b. \tag{5}$$

It follows that terms may be transferred from one side of a congruence to the other by a change of sign and also that congruences may be subtracted and the result is a valid congruence.

2) Congruences with respect to the same modulus may be multiplied and the result is a valid congruence.

$$\prod_{i=1}^{n} A_i \equiv \left(\prod_{i=1}^{n} \alpha_i \right) \text{Mod } b. \tag{6}$$

It follows that both sides of the congruence may be raised to the same power or multiplied by a constant and the result is a valid congruence.

3) Congruences are transitive. If $A \equiv B$ and $B \equiv C$, then $A \equiv C$.

4) A valid congruence relationship is obtained if the number, the residue and the modulus are divided by a common factor.

5) A valid congruence relationship is obtained if the number and the residue are divided by some common factor relatively prime to the modulus.

The material of this section has presented briefly, without proof, the pertinent concepts of congruences. Additional material on the subject may be found in any standard text on number theory.[1]

* Manuscript received by the PGEC, January 20, 1959; revised manuscript received March 31, 1959. An earlier version of this paper was presented at the 1959 WJCC, San Francisco, Calif.

† Dept. Electrical Engineering, University of Michigan, Ann Arbor, Mich.

[1] G. H. Hardy and E. M. Wright, "An Introduction to the Theory of Numbers," Oxford University Press, London, Eng.; 1956.

Development of the Residue Code

A residue code associated with a particular natural number is formed from the least positive residues of the particular number with respect to different bases. The first requirement for an efficient residue number system is that the bases of the difference digits of the representation must be relatively prime. If a pair of bases are not relatively prime, the effect is the introduction of redundancy. The following example will illustrate this fact. Contrast the residues associated with bases of magnitude 2 and 6 against the residues associated with bases of magnitude 3 and 4. In the first case, the bases are not relatively prime while in the second case the bases are relatively prime. The residues associated with the bases of magnitude 2 and 6 are unique for only 6 states while the residues associated with the bases of magnitude 3 and 4 provide a unique residue representation for 12 states. This is further clarified by Table I.

TABLE I

Redundancy of a Nonrelatively-Primed Base Representation

Number	Least Postive Residue			
	Mod 2	Mod 6	Mod 3	Mod 4
0	0	0	0	0
1	1	1	1	1
2	0	2	2	2
3	1	3	0	3
4	0	4	1	0
5	1	5	2	1
6	0	0	0	2
7	1	1	1	3
8	0	2	2	0
9	1	3	0	1
10	0	4	1	2
11	1	5	2	3
12	0	0	0	0
13	1	1	1	1
14	0	2	2	2

An example of a residue number system is presented in Table II. The number system shown in Table II uses the prime bases 2, 3, 5 and 7. The number system therefore contains 210 states. The 210 states may correspond to the positive integers 0 to 209. Table II shows the residue number representation corresponding to the positive integers 0 to 29. Additional integers of the number system may be found by congruence operations. Let a, b, c and d be the digits associated with the bases 2, 3, 5 and 7, respectively. The following congruences define a, b, c and d for the residue representation of the number N:

$$\begin{aligned} N &\equiv a \text{ Mod } 2 \\ N &\equiv b \text{ Mod } 3 \\ N &\equiv c \text{ Mod } 5 \\ N &\equiv d \text{ Mod } 7. \end{aligned} \tag{7}$$

The residue number system is readily extended to include more states. For example, if a base 11 is added to the representation, it is then possible to represent 2310 states. Table III shows the product and sum of the first nine consecutive primes greater than or equal to 2. The product of the primes indicates the number of states of the number system, while the sum of the primes is a measure of the size of the representation in terms of digits. Table III also includes the number of bits required to represent each prime base in the binary number system.

TABLE II

Natural Numbers and Corresponding Residue Numbers

N.N.	2357	N.N.	2357	N.N.	2357
0	0000	10	0103	20	0206
1	1111	11	1214	21	1010
2	0222	12	0025	22	0121
3	1033	13	1136	23	1232
4	0144	14	0240	24	0043
5	1205	15	1001	25	1104
6	0016	16	0112	26	0215
7	1120	17	1223	27	1026
8	0231	18	0034	28	0130
9	1042	19	1145	29	1241

TABLE III

Number of States and Digits Associated with a Residue Representation

i	p_i	$\sum_{i=1}^{n} p_i$	$\prod_{i=1}^{n} p_i$	p_i bits	$\sum p_i$ bits
1	2	2	2	1	1
2	3	5	6	2	3
3	5	10	30	3	6
4	7	17	210	3	9
5	11	28	2,310	4	13
6	13	41	30,030	4	17
7	17	58	510,510	5	22
8	19	77	9,699,690	5	27
9	23	100	223,092,670	5	32

Residue Addition and Multiplication

The residue number representation consists of several digits and is assumed to be in one to one correspondence with some positive integers of the real number system. The digits of the residue representation are the least positive residues of these real positive integers with respect to the different moduli which form the bases of the residue representation. It follows as a direct consequence of the structure of the residue number system and the properties of linear congruences that operations of addition and multiplication are valid in the residue number system subject to one proviso. The proviso is that the residue system must possess a number of states sufficient to represent the generated sum or product. If the residue number system does not have a sufficient number of states to represent the sums and the products generated by particular finite set of real integers then the residue system will overflow and more than one sum

or product of the real number system may correspond to one residue representation. For a residue number with a sufficient number of states an isomorphic relation exists with respect to the operations of addition and multiplication in the residue system and a finite system of real positive integers.

Each digit of the residue number system is obtained with respect to a different base or modulus. It follows, therefore, that the rules of arithmetic associated with each digit will be different. For example, the addition and multiplication of the digits associated with moduli 2 and 3 follow rules specified in Table IV. No carry

TABLE IV

MOD 2 AND MOD 3 SUMS AND PRODUCTS

$\oplus$	0	1
0	0	1
1	1	0

sum Mod 2

$\odot$	0	1
0	0	0
1	0	1

product Mod 2

$\oplus$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

sum Mod 3

$\odot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

product Mod 3

tables are necessary since the residue number system does not have a carry mechanism. Addition of two residue representations is effected by the modulo addition of corresponding digits of the two representations. Corresponding digits must have the same base or modulus. Modulo addition of digits which have different bases is not defined. Multiplication in the residue system is effected by obtaining the modulo product of corresponding digits. The operations of addition and multiplication of two residue numbers are indicated by the following notation:

$$S = A \oplus B$$
$$p = A \odot B. \tag{8}$$

Consider a residue number representation with bases 2, 3, 5 and 7. We assume an isomorphic relation between the residue number system and the real positive numbers 0 to 209. An isomorphic relation then exists for the operations of multiplication and addition only if the product or sum is less than 210. The following examples employing residue numbers illustrate the addition and multiplication operations and the presence of an isomorphism or the lack of isomorphism in the case of overflow. Residue numbers will be distinguished by the use of parentheses.

$$29 + 27 = S = 56$$
$$29 \leftrightarrow (1\ 2\ 4\ 1)$$
$$27 \leftrightarrow (1\ 0\ 2\ 6)$$
$$56 \leftrightarrow (0\ 2\ 1\ 0)$$
$$\begin{array}{r} (1\ 2\ 4\ 1) \\ \oplus\ (1\ 0\ 2\ 6) \\ \hline (0\ 2\ 1\ 0) \end{array}. \tag{9}$$

The following operations are considered in performing the addition of the two residue representations.

$$1 + 1 \equiv 0 \text{ Mod } 2$$
$$2 + 0 \equiv 2 \text{ Mod } 3$$
$$4 + 2 \equiv 1 \text{ Mod } 5$$
$$1 + 6 \equiv 0 \text{ Mod } 7. \tag{10}$$

Consider the addition of two numbers which produce a sum greater than 209.

$$S = 100 + 200$$
$$\begin{array}{r} (0\ 1\ 0\ 2) \\ \oplus\ (0\ 2\ 0\ 4) \\ \hline (0\ 0\ 0\ 6) \end{array}. \tag{11}$$

The residue representation (0 0 0 6) corresponds to the real positive number 90. In this particular example, the sum has overflowed the residue representation. The resulting sum is the correct sum modulo 210.

$$300 \equiv 90 \text{ Modulo } 210.$$

Finite real number systems and residue number systems have the same overflow characteristics. The sum which remains after the overflow is the correct sum with respect to a modulus numerically equal to the number of states in the finite number system.

The following is presented as an example of the process of residue multiplication.

$$p = 10 \times 17 = 170$$
$$10 \leftrightarrow (0\ 1\ 0\ 3)$$
$$17 \leftrightarrow (1\ 2\ 2\ 3)$$
$$170 \leftrightarrow (0\ 2\ 0\ 2)$$
$$\begin{array}{r} (0\ 1\ 0\ 3) \\ \odot\ (1\ 2\ 2\ 3) \\ \hline (0\ 2\ 0\ 2) \end{array}. \tag{12}$$

The process of multiplication involved consideration of the following relations for each digit.

$$1 \times 0 \equiv 0 \text{ Mod } 2$$
$$1 \times 2 \equiv 2 \text{ Mod } 3$$
$$0 \times 2 \equiv 0 \text{ Mod } 5$$
$$3 \times 3 \equiv 2 \text{ Mod } 7. \quad (13)$$

An overflow resulting from a multiplication is no different than the overflow resulting from an addition. Consider the product obtained from the residue multiplication of the numbers 10 and 100. The result in the modulo 210 number system is 160 since

$$1000 \equiv 160 \text{ mod } 210. \quad (14)$$

Subtraction and the Representation of Negative Numbers

The process of subtraction is obtainable in the residue number system by employing a complement representation consisting of the additive inverses of the positive residue representation. The additive inverse always exists, since each of the elements of the residue representation is an element of a field. There is no basic problem associated with the subtraction operation. There is, however, a problem associated with the representation of negative numbers. In particular, some mechanism must be included in the number system which will permit the representation of positive and negative numbers. This problem is discussed in this and the following section.

The additive inverse of a residue number is defined by the following:

$$a \oplus a' = 0. \quad (15)$$

The formula may be considered to apply to a digit of the residue system or equally well to the whole residue representation. Consider the following examples with reference to the modulo 210 residue number system.

$$a = (1\ 2\ 4\ 1) \quad (16)$$

then

$$a' = (1\ 1\ 1\ 6)$$

since

$$\begin{array}{r} (1\ 2\ 4\ 1) \\ \oplus\ (1\ 1\ 1\ 6) \\ \hline (0\ 0\ 0\ 0) \end{array}.$$

The following examples have been chosen to illustrate the subtraction process and to some extent the difficulties associated with the sign of the difference.

$$D = A \ominus B = A \oplus B'. \quad (17)$$

We consider first the case where the magnitude of A is greater than B.

Let

$$A = 200 \qquad B = 100.$$

In residue representation

$$B' = (0\ 2\ 0\ 5) \quad (18)$$

and

$$\begin{array}{r} (0\ 2\ 0\ 4) \\ \oplus\ (0\ 2\ 0\ 5) \\ \hline (0\ 1\ 0\ 2) \end{array}.$$

The residue representation of the difference corresponds to positive 100 in the real number domain. We consider next the difference

$$D = B \ominus A = A' \oplus B \quad (19)$$

where the magnitude of A is greater than B. Then

$$A' = (0\ 1\ 0\ 3) \quad (20)$$

and

$$\begin{array}{r} (0\ 1\ 0\ 3) \\ \oplus\ (0\ 1\ 0\ 2) \\ \hline (0\ 2\ 0\ 5) \end{array}.$$

The difference (0 2 0 5) is the additive inverse of (0 1 0 2). Unless additional information is supplied, the correct interpretation of the representation (0 2 0 5) is in doubt. (0 2 0 5) may correspond to either +110 or −100.

The difficulties associated with whether a residue representation corresponds to a positive or negative integer can be partially removed by the division of the residue number range into two parts. This is exactly the scheme that is employed to obtain a machine representation of positive and negative natural numbers. For the system of natural numbers two different machine representations of the negative numbers may be obtained and are commonly designated the radix complement representation of negative numbers and the diminished radix complement representation of negative numbers.

The complement representation for a residue code is defined in terms of the additive inverse. Thus, the representation of negative A is A' where $A \oplus A' = 0$, and the range of A is restricted to approximately one-half of the total possible range of the residue representation. This can be illustrated by consideration of a specific residue code. This residue representation, employing bases of magnitude 2, 3, 5, and 7, is divided into two parts. The residue representations corresponding to the natural numbers 0 to 104 are considered positive. The residue representations corresponding to the natural numbers 105 to 209 are considered inverse representations and associated with the negative integers from −1 to −105.

The range of this particular number system is from -105 to $+104$. The arithmetic rules pertaining to sign and overflow conventions for this particular number system are the same rules normally associated with radix complement arithmetic.

The complement representation does eliminate in principle any ambiguity concerning the sign of the result of an arithmetic operation. However, there is a practical difficulty. The determination of the sign associated with a particular residue representation requires the establishment of the magnitude of the representation relative to the magnitude which separates the positive and negative representations. The determination of relative magnitude for a residue representation is discussed in the next section. It will be shown that the determination of relative magnitude is not a simple problem.

Conversion from a Residue Code to a Normal Number Representation

It is frequently desirable to determine the natural number associated with a particular residue representation. The need for this conversion occurs frequently when investigating the properties of the residue system. The residue representation is constructed in such a manner that magnitude is not readily obtainable. The presence of digit weights in the normal polynomial type number representation greatly facilitates the determination of magnitude. However, it is possible to assign a weight to each digit of the residue representation in such a manner that the modulo m sum of the digit-weight products is the real natural number in a consistently weighted representation. m is the product of all the bases employed in the residue representation. The conversion technique is known as "The Chinese Remainder Theorem." The material which follows describes the conversion technique but omits the proof. A simple and straightforward proof is found in Dickson.[2] The proof does not refer specifically to residue number systems, but rather to a system of linear congruences. If so regarded, a system of congruences defines a component of a residue number system.

Consider a residue number system with bases $m_1 \cdots m_t$. The corresponding digits are labeled $a_1 \cdots a_t$. The following equations define the conversion process.

$$a_1 A_1 \frac{m}{m_1} + \cdots + a_t A_t \frac{m}{m_t} \equiv S \text{ Mod } m \qquad (21)$$

where

$$A_i \frac{m}{m_i} \equiv 1 \text{ Mod } m_i$$

and

$$m = \prod_{j=1}^{t} m_j.$$

The conversion formula for a particular residue number system is now obtained.

$$m_1 = 2 \quad m_2 = 3 \quad m_3 = 5 \quad m_4 = 7$$

$$105\, A_1 \equiv 1 \text{ Mod } 2 \quad \text{so } A_1 = 1$$

$$70\, A_2 \equiv 1 \text{ Mod } 3 \quad \text{so } A_2 = 1$$

$$42\, A_3 \equiv 1 \text{ Mod } 5$$

$$2\, A_3 \equiv 1 \text{ Mod } 5 \quad \text{so } A_3 = 3$$

$$30\, A_4 \equiv 1 \text{ Mod } 7$$

$$2\, A_4 \equiv 1 \text{ Mod } 7 \quad \text{so } A_4 = 4$$

$$105\, a_1 + 70\, a_2 + 126\, a_3 + 120\, a_4 \equiv S \text{ Mod } 210. \qquad (22)$$

The conversion formula is now used to determine the natural number corresponding to the residue representation (1 2 0 4).

$$105\,(1) + 70\,(2) + 126\,(0) + 120\,(4) = 725$$

$$725 \equiv S \text{ Mod } 210$$

$$S = 95. \qquad (23)$$

The above conversion formula may be obtained in a much more direct manner. The residue number representation may be regarded as a vector with the base vectors (1000), (0100), (0010) and (0001). The residue representation expressed in terms of the base vectors is:

$$(a_1 a_2 a_3 a_4) = a_1(1000) \oplus a_2(0100) \oplus a_3(0010) \oplus a_4(0001). \qquad (24)$$

The magnitude of the base vectors is

$$(1000) \leftrightarrow 105$$

$$(0100) \leftrightarrow 70$$

$$(0010) \leftrightarrow 126$$

$$(0001) \leftrightarrow 120. \qquad (25)$$

Other conversion techniques exist. In particular it is possible by means of a deductive process to determine the magnitude of a particular residue representation. This requires both a knowledge of the nature of the residue system and the natural number representation associated with at least one residue representation.

Due to the deductive nature of the process, it is more suitable for human computation than for machine computation. The process is explained using the residue number of the previous example (1 2 0 4). The knowledge of the residue representation for unity which is (1 1 1 1) is assumed. Consider the effect of changing the second digit from one to two. The change adds the product $m_1 m_3 m_4 = 70$ to the number since 70 is congruent 1, modulo 3. The resulting residue representation (1 2 1 1) corresponds to 71. The effect of changing the third digit is to change the magnitude by some multiple of the product $m_1 m_2 m_4 = 42$. The correct change in magnitude is $42x$ where $42x \equiv 4$ Mod 5. So $42x = 84$ and the

[2] L. E. Dickson, "Modern Elementary Theory of Numbers," University of Chicago Press, Chicago, Ill., p. 16: 1939.

residue representation (1 2 0 1) corresponds to 155. The fourth digit is modified by the addition of a three. The effect of this change is determined by $30x \equiv 3 \text{ Mod } 7$. The magnitude change is 150. The sum of 150 and 155 modulo 210 yields the correct result 95, in correspondence with (1 2 0 4).

Sign determination for the residue code is dependent on the determination of a greater than or less than relationship. A possible method might involve the conversion techniques described previously. Such a scheme would involve the standard comparison techniques associated with the determination of the relative magnitude of two numbers represented in a weighted representation. An alternate conversion procedure yields a conversion from the residue code to a nonconsistently based polymonial number representation by means of residue arithmetic. Consider a residue code consisting of t digits. The t digits of the residue code are associated with t congruence relationships as follows:

$$S \equiv a_i \text{ Mod } m_i \qquad 1 \leq i \leq t \tag{26}$$

S is the magnitude of the number expressed in normal representation. It is also possible to express the number S as

$$S = a_i + A_i m_i. \tag{27}$$

A_i is the integer part of the quotient of S divided by m_i. In regard to a greater or less than relationship, the determination of A_i divides the range of the residue representation into m/m_i parts. We proceed to calculate A_i from the set of t equations given above. Let

$$S = a_t + A_t m_t \qquad A_t < \frac{m}{m_t}. \tag{28}$$

This equation is then used to replace S in the remaining $t-1$ equations, yielding $t-1$ equations of the form

$$A_t m_t \equiv (a_i + a_t') \text{ Mod } m_i \quad 1 \leq i \leq t-1 \tag{29}$$

or

$$A_t \equiv (a_i + a_t')/m_t^i \text{ Mod } m_i$$

$$A_t \equiv d_t^i \text{ Mod } m_i$$

where $/m_t^i$ is the multiplicative inverse of m_t with respect to base m_i. The multiplicative inverse is defined as[3]

$$x_t/x_t^i \equiv 1 \text{ Mod } m_i \tag{30}$$

d_t^i is the least positive residue of $(a_i + a_t')/m_t$ with respect to base i. a_t' is the additive inverse of a_t. Let A_t be expressed as

$$A_t = d_t^{t-1} + A_{t-1} m_{t-1} \qquad A_{t-1} < \frac{m}{m_t m_{t-1}}. \tag{31}$$

[3] The existence of the multiplicative inverse requires that x_t and m_i be relatively prime.

If this expression is substituted for A_t a set of $t-2$ equations remain. The equations are of the form

$$A_{t-1} \equiv [d_t^i + (d_t^{t-1})']/m_{t-1} \text{ Mod } m_i \quad 1 \leq i \leq t-2$$

$$A_{t-1} \equiv d_{t-1}^i \text{ Mod } m_i. \tag{32}$$

The system of equations shown below is generated by repetition of the above substitution process until no equations remain.

$$S = a_t + A_t m_t \tag{33}$$

$$A_t = d_t^{t-1} + A_{t-1} m_{t-1}$$

$$A_{t-1} = d_{t-1}^{t-2} + A_{t-2} m_{t-2}$$

$$\vdots$$

$$A_3 = d_3^2 + A_2 m_2$$

$$A_2 \equiv d_2^1 \text{ Mod } m_1. \tag{34}$$

The equations are combined to yield:

$$S = a_t + m_t \{ d_t^{t-1} + m_{t-1}[d_{t-1}^{t-2} + m_{t-2}(d_{t-2}^{t-3} + \cdots \tag{35}$$

$$= a_t + m_t d_t^{t-1} + m_t m_{t-1} d_{t-1}^{t-2}$$

$$+ m_t m_{t-1} m_{t-2} d_{t-2}^{t-3} + \cdots + \frac{m}{m_1} d_2^1 \tag{36}$$

where

$$A_t < m_t$$

$$d_{t-n}^{t-n-1} < m_{t-n-1}.$$

Therefore, S is never equal to or greater than m and d_2^1 divides the range into m_1 parts, d_3^2 divides each of the m_1 parts into m_2 parts, d_4^3 divides each of the m_2 parts into m_3 parts, etc.

The formulas which define the conversion process may be applied recursively to obtain a formula for a greater number of digits. The process has been extended to five variables and the results are shown in Fig. 1.

A somewhat simpler interpretation of the conversion process is obtained if (36) is considered from the vector viewpoint. Let the terms $1, m_t, m_t m_{t-1}, \cdots, m/m_1$ be regarded as base vectors. The residue representation of the base vectors is readily obtainable and the following characteristics with respect to zero and nonzero digits of the base vectors are observed

$$1 \leftrightarrow (xxx \cdots xxx)$$

$$m_t \leftrightarrow (xxx \cdots xx0)$$

$$m_t m_{t-1} \rightarrow (xxx \cdots x00) \quad x \text{ indicates a nonzero digit}$$

$$\vdots \qquad\qquad\qquad\qquad 0 \text{ indicates a zero digit}$$

$$\frac{m}{m_1} \leftrightarrow (x00 \cdots 000). \tag{37}$$

The following equality exists between the vector representation of the residue code and the vector representation of the nonconsistently based code.

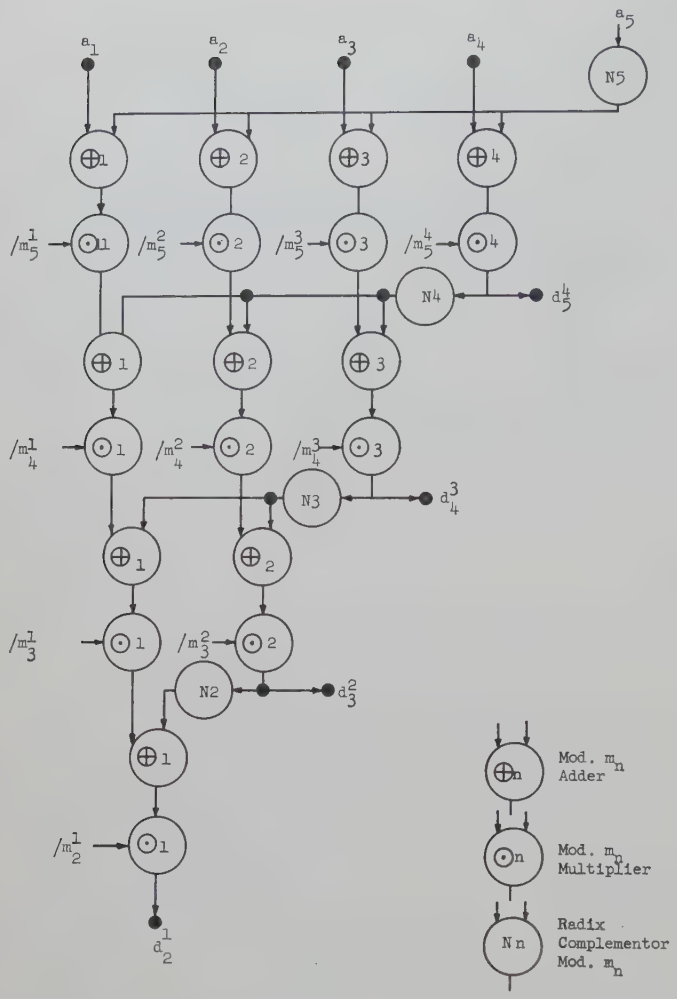


Fig. 1—Logic required to convert from the residue code to a weighted code.

$$a_1(101\cdots000)\oplus a_2(010\cdots000)\oplus\cdots$$
$$\oplus a_{t-1}(000\cdots010)\oplus a_t(000\cdots001)$$
$$= a_t(111\cdots111)\oplus d_t^{t-1}(1xx\cdots xx0)\oplus\cdots$$
$$\oplus d_3^2(1x0\cdots000)\oplus d_2^1(100\cdots000). \tag{38}$$

Equating the coefficients of each coordinate in (38) obtains the following set of equations relating the residue digits and the digit of the nonconsistently based number system.

$$\begin{aligned} a_t &\equiv a_t \bmod m_t \\ xd_t^{t-1} + a_t &\equiv a_{t-1} \bmod m_{t-1} \\ &\vdots \\ xd_3^2 + \cdots + xd_t^{t-1} + a_t &\equiv a_2 \bmod m_2 \\ d_2^1 + d_3^2 + \cdots + d_t^{t-1} + a_t &\equiv a_1 \bmod m_1. \end{aligned} \tag{39}$$

The conversion is executed by first subtracting $a_t(111\cdots111)$; $d_t{}^{t-1}$ is then a function of only one residue digit and may be determined. This is followed by the subtraction of $d_t{}^{t-1}(1xx\cdots xx0)$ etc. Consider the conversion of the residue representation (1204) with bases 2, 3, 5 and 7.

The weights and the corresponding residue representation of the nonconsistently based code are:

$$\begin{aligned} 1 &\leftrightarrow (1111) \\ m_4 = 7 &\leftrightarrow (1120) \\ m_4m_3 = 35 &\leftrightarrow (1200) \\ m_4m_3m_2 = 105 &\leftrightarrow (1000). \end{aligned} \tag{40}$$

The actual conversion between the residue code and the nonconsistently based code is accomplished by the following operation:

	(120④)	1st digit is a 4 since $4\cdot1\equiv4 \bmod 7$
$\oplus$	(0213)	Subtract off 4(1111) = (0144)
	11①0	2nd digit is a 3 since $3\cdot2\equiv1 \bmod 5$
$\oplus$	104	Subtract off 3(1120) = (1010)
	0①00	3rd digit is a 2 since $2\cdot2\equiv1 \bmod 3$
$\oplus$	02	Subtract off 2(1200) = (0100)
	00	4th digit is 0.

Thus

$$\begin{aligned} S &= 105\cdot0 + 35\cdot2 + 7\cdot3 + 1\cdot4 \\ &= 95. \end{aligned} \tag{41}$$

Admittedly, the process required to obtain a magnitude leaves much to be desired though the conversion process is no more complicated than the standard change of base conversion process used for consistently based number systems. One presumed advantage of the residue number system was the absence of a carry process. The greater or less than process is essentially sequential and is in many ways similar to the carry process of ordinary arithmetic. The ultimate usefulness of the residue code for general purpose computation appears very much dependent on the development of simple techniques for the determination of the relative magnitude of two residue code digits.

Division

The division process for residue codes is complicated by two factors. The first is the absence of a multiplicative inverse for the zero element. The second difficulty is the fact that residue division and the normal division process are in one to one correspondence only when the resulting quotient is an integer value. We shall consider first the problem of residue division of the elements of a single field and shall consider later the elements of several fields considered as a residue code. The division process represented in equation form as

$$\frac{a}{b} = q \tag{42}$$

implies the following equation:

$$a = bq. \tag{43}$$

The difference between normal arithmetic and residue arithmetic is that in residue arithmetic the product bq

need not necessarily be equal to a, only the congruence of a and bq is required.

$$bq \equiv a \text{ Mod } m_n. \tag{44}$$

Multiplication by the multiplicative inverse of b designated $/b$ obtains

$$q \equiv a/b \text{ Mod } m_n. \tag{45}$$

The correct interpretation of q in the above equation is that the number a is obtained by forming the modulo sum consisting of b, q representations. The sum is carried out in a closed and finite modulo number system of base m_n. Thus, q corresponds to the quotient only when the quotient has an integer value. Examples may be obtained from the consideration of a modulo 5 number system

$$\frac{4}{2} = q \tag{46}$$

$$2q \equiv 4 \text{ Mod } 5$$

$$q \equiv 2 \text{ Mod } 5$$

$$\frac{4}{3} = q \tag{47}$$

$$3q \equiv 4 \text{ Mod } 5$$

$$q \equiv 3 \text{ Mod } 5$$

$$\text{note } 3 \times 3 \equiv 4 \text{ Mod } 5$$

$$\frac{3}{4} = q \tag{48}$$

$$4q \equiv 3 \text{ Mod } 5$$

$$q \equiv 2 \text{ Mod } 5.$$

In the above examples, q corresponds to the quotient only in the first example.

The residue code representation of a number consists of many digits, $A = (a_1, a_2, \cdots, a_n)$. Each digit of the representation is associated with a different prime base. The number system is a modulo m system where

$$m = \prod_{i=1}^{n} m_i.$$

The division of two numbers in residue code may be expressed by a system of congruences. The solution $Q = (q_1, q_2, \cdots, q_n)$ must satisfy all the congruence relationships of the system. A zero digit in the divisor $B = (b_1, b_2, \cdots, b_n)$ means that B and m are not relatively prime hence the multiplicative inverse of B doesn't exist.

$$QB \not\equiv A \text{ Mod } m. \tag{49}$$

For the special case in which $b_i = 0$ and $a_i = 0$, a valid congruence relationship of the form

$$\frac{QB}{m_i} \equiv \frac{A}{m_i} \text{ Mod } \frac{m}{m_i} \tag{50}$$

is obtainable.

The process of residue division has certain interesting properties and quite possibly has applications in respect to special problems. Unfortunately, the residue division process is not a substitute for normal division. It appears that the only way in which division can be effected in the residue code is by the utilization of techniques similar to those employed for division in a consistently weighted number system. The division process then requires trial and error subtraction or addition and the greater than or less than relationship. The division algorithm could also include trial multiplication since in the residue system addition and multiplication require the same period of time.

Conclusions

The material of this paper forms a preliminary investigation of the applicability of residue number systems to the arithmetic operations of digital computers. The residue system has been found attractive in terms of the operations of multiplication and addition. It is possible to realize practical logical circuitry to yield the product in the same operation time as for the sum since the product is not obtained by the usual procedure of repetitive addition. The main disadvantages of the residue number system are associated with the necessity of determining absolute magnitude. Thus, the division process, the detection of an overflow and the determination of the correct sign of a subtraction operation are processes which at this stage of the investigation seem to involve considerable complexity. Nevertheless, there are certainly many special purpose applications well-suited to the residue code. In particular, there exists a class of control problems characterized by the absence of the need for division, the existence of a well-defined range for the variables and also by the fact that the sign of the variables is known. For the problems of this class, the use of the residue code should result in a reduction of the over-all computation period and give a computer with a higher bandwidth than obtainable with the conventional number system.

The ultimate usefulness of the residue code will probably be determined largely by the success of the circuit designer in perfecting circuitry ideally suited for residue code operations.

Acknowledgment

The material of this paper is essentially Chapter 5 of the author's Ph.D. dissertation.[4] At the time of the completion of the dissertation, the author was unaware of the work of Valach[5] and Svoboda[6,7] in Czechoslovakia, references of which were obtained from recent visitors to the Soviet Union. The author wishes to take this opportunity to acknowledge same.

[4] H. L. Garner, "Error Checking and the Structure of Binary Addition," Ph.D., dissertation, University of Michigan, Ann Arbor, Mich., pp. 105–140; 1958.

[5] M. Valach, "Vznik kodu a ciselne soustavy zbytkovych trid," *Stroje Na Zpracovani Informaci*, Sbornik III; 1955.

[6] A. Svoboda and M. Valach, "Operatorove Obvody," *Stroje Na Zpracovany Informaci*, Sbornik III; 1955.

[7] A. Svoboda, "Rational Numerical System of Residual Classes," *Stroje Na Zpracovani Informaci*, Sbornik V; 1957.

Bibliography of Digital Magnetic Circuits and Materials*

WALTER L. MORGAN†

Summary—A list of about 400 references relating to magnetic memory and logic circuits has been made. The bibliography is divided into 19 sections. Several sections are devoted to the physical, magnetic, and switching parameters of magnetic materials. Other parts cover the circuit and logical aspects of using magnetic cores, plates, "twistors," thin films, and transfluxors. Attention is given to the use of special memory techniques such as domain wall viscosity readout, cross-field effects, and circuits operated with RF carriers. The use of magnetic cores as half adders, gates, and shift registers is recognized in a separate section. A listing of sources of further information (conference proceedings, books, and other bibliographies) is included.

Table of Contents

Introduction

THE magnetic materials and circuits covered in this bibliography are restricted to those used in digital circuits involving no moving parts. Other magnetic storage methods (such as recording on wire, tape, or drum) have not been included; these have been previously treated in other lists.[1] Another area that is not listed is that of the patent field.

Most of the references are available from any large corporate, educational, or institutional library. In many instances a paper may have been published in several different journals; duplicate listings have been deliberately retained to make the selected article easier to find.

Although there is a vast amount of literature available concerning the physics and metallurgy of magnetic-core materials, only a representative cross section is included. This should be enough to give a logical circuit designer a good physical background without overwhelming him in details. There are many other fine references which, unfortunately, had to be omitted because of space limitations. Many of these may be found by consulting section 18, Other Bibliographies.

Magnetic circuits have been used primarily for the storage of digital information. While an attempt has been made to separate the driving circuits from the memory array fabrication and operation, there are many single articles that deal with both; therefore, it is suggested that literature dealing with peripheral equipment be examined, in addition to the memory references, by someone seeking sources of information about magnetic memory planes (and vice versa).

Several sections are devoted to special types of magnetic components. These include flux-logic (the "transfluxor") and the "twistor."

Magnetic shift registers have long been used in digital machines for the temporary storage of information, speed conversion, and serial-parallel (or parallel-serial) converters. Magnetic logic circuits have been described by many authors. A list of these has been included.

Journal abbreviations follow those used in the Abstract and Reference Sections of *Electronics and Radio Engineering* (incorporating *Wireless Engineer*) and the PROCEEDINGS OF THE INSTITUTE OF RADIO ENGINEERS. The abbreviation vol. 51.1, *e.g.*, refers to vol. 51, no. 1, and is used primarily in those cases where each issue has page numbers beginning from 1.

1. Magnetic Materials

PART A. Magnetic Metals (4-79 Mo-Permalloy, Etc.)

M. F. Littman, "Ultrathin magnetic alloy tapes with rectangular hysteresis loops," *Elec. Engrg.*, vol. 71, p. 792; September, 1952.

"Making thin-gauge magnetic cores," *Elec. Mfg.*, vol. 51.1, pp. 140–141; March, 1953.

M. Kornetzki, J. Brackmann, and J. Frey, "Ferritkerne mit Perminvar-Schleife (Ferrite cores with perminvar loop)," *Siemens Z.*, vol. 29, pp. 434–440; 1955. (In German.)

G. H. Howe, "Dynamax, a new crystal and domain oriented magnetic core material," *Commun. & Electronics*, no. 27 (*Trans. AIEE*, pt. I, vol. 75), pp. 548–550; November, 1956.

A. L. Freedman, "Magnetic core coincidence circuit," *Electronics*, vol. 30, pp. 194–196; August 1, 1957.

R. S. Sery, R. E. Fischell, and D. I. Gordon, "Radiation Effects on Magnetic Materials," Naval Ordinance Lab. Rept. 4381; abstracted in *Electronics*, vol. 30, p. 250; October 1, 1957.

W. Arrott, "Design advances from magnetism research," *Elec. Mfg.*; January, 1958.

M. A. Rozenblat, "Dynamic characteristics of cores with a rectangular static hysteresis loop (The effect of eddy currents)," *Automatika i Telemekhanika*; January, 1958. (In Russian.)

J. A. Osborn, R. C. Hill, J. F. Fritz, J. J. Clark, *et al.*, "Solid State Research for the Advancement of Electronic Materials," Westing-

* Manuscript received by the PGEC, December 1, 1958; revised manuscript received, March 27, 1959.

† RCA, Moorestown, N. J.

[1] See, for instance, C. F. Wilson, "Magnetic recording, 1888–1952," IRE TRANS. ON AUDIO, vol. AU-4, pp. 53–81; May/June, 1956.

house Elec. Corp.; WADC TR 58-21, AF 33 (616)309; January, 1958. Available as ASTIA AD 142293.
W. D. McMiller and D. Howell, "The Effect of Reactor Radiation on the Electrical Properties of Electronic Components, Part IV; Capacitors and Magnetic Cores," Convair, Fort Worth, Tex., NARF-58-6T, MR-N-185, AF 33(600)32054; January 28, 1958.
D. I. Gordon, R. S. Sery, and R. E. Fischell, "Radiation effects in magnetic materials," *Nucleonics*, vol. 16.6, pp. 73–77; June, 1958.

PART B. Ferrites

C. L. Snyder, E. Albers-Schoenberg, and H. Goldsmith, "Magnetic ferrites," *Elec. Mfg.*, vol. 44, p. 86; 1949.
D. Polder, "Ferrite materials," *Proc. IEE*, vol. 97, pt. 2, pp. 246–256; April, 1950.
M. J. O. Strutt, "Ferromagnetic materials and ferrites," *Wireless Engr.*, vol. 27, pp. 277–284; December, 1950.
B. V. Vonderschmitt, M. J. Obert, and H. B. Stott, "Ferrite applications in electronic components," *Electronics*, vol. 25.3, pp. 138–139; March, 1952.
E. Both, "Development and utilization of magnetic ferrites," *Ceramic Age*, vol. 59.4, pp. 39–45; April, 1952.
E. Emmerich and S. C. Klaus, "Evaluation of electrical properties of ferrite materials for specific applications," *Ceramic Age*, vol. 59.5, pp. 26–29; May, 1952.
E. Gelbard, "Magnetic properties of ferrite materials," *Tele-Tech.*, vol. 11; May, 1952.
M. Kornetzki, J. Brackmann, J. Frey, and W. Gieseke, "Measurements on high permeability ferrite cores," *Z. angew Math. Phys.*, vol. 4; October, 1952 (in German).
C. D. Owens, "Analysis of measurements on magnetic ferrites," Proc. IRE, vol. 41, pp. 359–365; March, 1953.
R. M. Bozorth, "Behavior of magnetic materials," *Amer. J. Phys.*, vol. 21, pp. 260–266; April, 1953.
M. Kornetzki, "Magnetic constants of ferrites," *Siemens Entwicklungs Berichte*, vol. 16; July, 1953 (in German).
H. J. Williams, R. C. Sherwood, M. Goertz, and F. J. Schnettler, "Stressed ferrites having rectangular hysteresis loops," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 72), pp. 531–537; November, 1953.
I. J. Hegyi, "Ferromagnetic spinels with rectangular hysteresis loops," *J. Appl. Phys.*, vol. 25, pp. 176–178; February, 1954.
E. Albers-Schoenberg, "Ferrites for microwave circuits and digital computers," *J. Appl. Phys.*, vol. 25, pp. 152–154; February, 1954.
H. P. J. Wijn, E. W. Gorter, C. J. Esvaldt, and P. Gildermans, "Conditions for square hysteresis loops in ferrites," *Philips Tech. Rev.*, vol. 16, pp. 49–58; August, 1954.
V. E. Legg and C. D. Owens, "Magnetic ferrites—new materials for modern applications," *Elec. Engrg.*, vol. 73; August, 1954.
O. Eckert, "Ferrite mit rechteckfoermiger Hystereseschleife," *Elektrotech. Z.*, vol. A-75, p. 253; 1954.
E. Albers-Schoenberg and O. Eckert, "Die Keramischen 'Rechteck'-Ferrite," *Ber. d. D. K. G.*, vol. 31, p. 311; 1954.
D. R. Brown, D. A. Buck, and N. Menyuk, "A comparison of metals and ferrites for high speed pulse operation," *Commun. & Electronics* (*Trans. AIEE*, vol. 73), pp. 631–635; (1954) January, 1955.
R. M. Bozorth, "Magnetic materials," *Scientific American*, vol. 192, pp. 68–73; January, 1955.
M. Kornetzki, "A note on the rectangular magnetization loop of ferrite cores," *Frequenz*, vol. 9, pp. 81–83; March, 1955.
G. S. Hipskind, "Losses in magnetic ferrites," *Elec. Mfg.*, pp. 108–112; August, 1955.
C. F. Devenny and L. G. Thompson, "Ferromagnetic computer cores," *Tele-Tech & Electronic Ind.*, vol. 14, pp. 58–59, 84–94; September, 1955.
O. Eckert, E. Weides, and K. Wallenfang, "Assessment of quality of rectangular (-loop) ferrites for electronic computers," *Elecktronische Rundschau*, vol. 9, pp. 371–374; October, 1955.
E. W. Gorter, "Some properties of ferrites in connection with their chemistry," Proc. IRE, vol. 43, pp. 1945–1973; December, 1955 (contains a list of 63 references).
N. I. Ananthanarayanan, "Soft magnetic materials with rectangular hysteresis loops," *Proc. AIEE Conf. on Magnetism and Magnetic Materials* (AIEE Publication T-78), pp. 236–246; 1955 (contains a list of 86 references).
V. C. Wilson, "Applications of ferrites," *Proc. AIEE Conf. on Magnetism and Magnetic Materials* (AIEE Publication T-78), pp. 278–285; 1955.
R. S. Duncan and H. A. Stone, Jr., "A survey of the application of ferrites to inductor design," Proc. IRE, vol. 44, pp. 4–13; January, 1956 (contains 38 references).
E. Albers-Schoenberg, "Square hysteresis loop ferrites in the system MgO—MnO—Fe_2O_3," *Amer. Ceramic Soc. Bull.*, vol. 35, pp. 276–278; July, 1956.
C. D. Owens, "Modern magnetic ferrites and their engineering applications," IRE Trans. on Component Parts, vol. CP-3, pp. 54–62; September, 1956.
C. D. Owens, "A survey of the properties and applications of ferrites below microwave frequencies," Proc. IRE, vol. 44, pp. 1234–1248; October, 1956 (contains 70 references).
"Progress in ferrite materials," *Electronics & Radio Engrg.*, vol. 34, pp. 56–61; February, 1957.
G. Campbell, "Some new electrical and magnetic ceramics," *J. Sci. Instr.*, vol. 34, pp. 337–348; September, 1957.
A. H. Iverson, "Hermetic seal for ferrites," *Rev. Sci. Instr.*, vol. 28, pp. 797–799; October, 1957.
V. E. Legg, "Survey of square loop magnetic materials," IRE Trans. on Component Parts, vol. CP-4, pp. 106–109; December, 1957.
E. W. Gorter and C. J. Esveldt, "Square loop ferrites obtained by magnetic annealing of new compositions," *Proc. IEE*, vol. 104, pt. B, supplement 7, pp. 418–421; 1957.
"Convention on ferrites," *Proc. IEE*, vol. 104, pt. B, supplements 5 and 7, pp. 127–266, 399–570; 1957 (a group of 40 papers dealing with the chemistry and physics of ferrites, and application of ferrite square-loop properties).
T. H. Bonn, R. D. Torrey, and F. Bernstein, "B-H tester measures memory core parameters," *Electronics*, vol. 31.2, pp. 76–80; January 17, 1958.
W. Arrott, "Design advances from magnetism research," *Elec. Mfg.*; January, 1958 (based on papers presented at the AIEE Third Natl. Conf. on Magnetism and Magnetic Materials).
A. D. Holt, "The application of square hysteresis loop materials in digital computer circuits," *Electronic Engrg.*; April, 1958.
C. Heck and H. Reiner, "Square loop ferrites and tests of their storage properties," *Nachrichtentech. Z.*; July, 1958.
E. Albers-Schoenberg, "Ferrites," *J. Amer. Ceramic Soc.*, vol. 41, pt. II, pp. 484–489; November, 1958.
S. L. Blum, "Microstructure and properties of ferrites," *J. Amer. Ceramic Soc.*, vol. 41, pt. II, pp. 489–493; November, 1958.
A. C. Hudson and E. J. Stevens, "Data on ferrite core materials," *Electronic Engrg.*, vol. 30, pp. 718–719; December, 1958.
S. A. Friedberg, "Heat conduction in some ferrimagnetic crystals at low temperatures," (Abstract only) *J. Appl. Phys.*, suppl. to vol. 30.4, p. 21S; April, 1959.
W. Heister, "Magnetic properties and granular structure of Mn-Zn-ferrites," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 22S–24S; April, 1959. *Proc. 4th Symp. on Magnetism and Magnetic Materials*, Philadelphia, Pa.; November 17–20, 1958, *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 1S–322S; April, 1959.

PART C. Thin Films

(*See also Section 7, Thin Film Memory Elements*)

M. S. Blois, Jr., "Evaporated magnetic materials," *Electronics*, vol. 28, pp. 210, 212, 214, 216, and 218; June, 1955.
M. S. Blois, Jr., "Preparation of thin films and their properties," *J. Appl. Phys.*, vol. 26, pp. 975–980; August, 1955.
D. A. Buck, "The cryotron, a superconductive computer component," Proc. IRE, vol. 44, pp. 482–493; April, 1956.
D. O. Smith, "Magnetic relaxation in thin films," *Proc. AIEE Conf. on Magnetism and Magnetic Materials*, Boston, Mass.; October, 1956.
R. L. Conger, "High frequency effects in magnetic films," *Proc. AIEE Conf. on Magnetism and Magnetic Materials*, Boston, Mass., p. 610; October, 1956.
H. B. Callen, "Rotational remagnetization," paper presented at the Internatl. Symp. on Switching Theory, Harvard University, Cambridge, Mass.; April 2–5, 1957.
R. L. Conger and F. C. Essig, "Magnetization reversal in thin films at low fields," *J. Appl. Phys.*, vol. 28, pp. 855–858; August, 1957.
J. W. Crowe, "Trapped-flux superconducting memory," *IBM J. Res. and Dev.*, vol. 1, pp. 294–303; October, 1957.
R. L. Garwin, "An analysis of a persistent-superconducting memory cell," *IBM J. Res. and Dev.*, vol. 1, pp. 304–308; October, 1957.
W. Arrott, "Design advances from magnetism research," *Elec. Mfg.*, pp. 86–100; January, 1958.

H. W. Katz, "Review of several solid-state devices and their application," paper presented at the 1958 Transistor and Solid-State Circuits Conf., Philadelphia, Pa.; February 20–21, 1958.
D. O. Smith, "Static and dynamic behavior of thin permalloy films," *J. Appl. Phys.*, vol. 29, pp. 264–273; March, 1958.
J. Kaczar, "Contributions to the theory of coercive force in thin sheets," *Bull. Acad. Sci. USSR*, vol. 21.8, pp. 1168–1176; 1958.
Proc. 4th Symp. on Magnetism and Magnetic Materials, Philadelphia, Pa.; November 17–20, 1958. *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 1S–322S; April, 1959.

2. Pulse Response of Magnetic Circuits

E. A. Sands, "Behavior of rectangular hysteresis loop magnetic materials under current pulse conditions," Proc. IRE, vol. 40, pp. 1246–1250; October, 1952.
J. K. Galt, J. Andrus, and H. G. Hooper, "Motion of domain walls in ferrite crystals," *Rev. Mod. Phys.*, vol. 25, pp. 93–97; January, 1953.
O. J. Van Sant, Jr. and E. D. Burnside, "A Method of Measuring Flux-Current Loops under Pulse Excitation," NAVORD Rept. 2970, U. S. Naval Ordnance Lab., White Oak, Md.; October, 1953.
J. Wylen, "Pulse response characteristics of rectangular hysteresis-loop ferromagnetic materials," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 72), pp. 648–655; November, 1953.
O. J. Van Sant, Jr., "Considerations for the selection of magnetic core materials for digital computer elements," 1954 IRE Convention Record, pt. 4, pp. 109–113.
R. C. Minnick, "Magnetic switching circuits," *J. Appl. Phys.*, vol. 25, p. 479; April, 1954.
J. R. Freeman, "Pulse Responses of Ferrite Memory Cores," M.I.T. Lincoln Lab., Lexington, Mass., Div. 6, Memor. M-2568-1. Also, Proc. 1954 WESCON Comp. Sessions, pp. 50–62.
D. S. Rodbell and C. P. Bean, "Influence of pulsed magnetic fields on the reversal of magnetization in square-loop metallic tape," *J. Appl. Phys.*, vol. 26, pp. 1318–1323; November, 1955.
D. S. Rodbell and C. P. Bean, "The influence of pulse magnetic fields on the reversal of magnetization in square-loop metallic tapes," *Proc. AIEE Conf. on Magnetism and Magnetic Devices—1955* (AIEE Publication T-78), pp. 219–223.
T. Douma, "Magnetization current in pulse transformers," *Proc. Natl. Electronics Conf.*, Chicago, Ill., vol. 12, pp. 1043–1061; 1956.
R. Kikuchi, "On the minimum of magnetization reversal time," *J. Appl. Phys.*, vol. 27, pp. 1352–1357; November, 1956.
J. D. Childress, "Effects of field pulse rise waveform on the pulse response of rectangular loop ferrites," *Proc. AIEE Solid-State Dielectric and Magnetic Devices Conf.*, Washington, D. C.; April 22–23, 1957.
R. W. Roberts and C. C. Horstman, "Core tester simplifies ferro-amplifier design," *Electronics*, vol. 30.8, pp. 150–153; August 1, 1957.
V. L. Newhouse, "The utilization of domain wall viscosity in data-handling devices," Proc. IRE, vol. 45, pp. 1484–1494; November, 1957.
M. A. Rozenblat, "Dynamic characteristics of cores with a rectangular static hysteresis loop (The effect of eddy currents)," *Automatika i Telemekhanika*; January, 1958. (In Russian.)
T. H. Bonn, R. D. Torrey, and F. Bernstein, "B-H tester measures memory core parameters," *Electronics*, vol. 31, pp. 76–80; January 17, 1958.
C. E. Goodell, "Integrator-amplifier for core measurements," *Electronics*, vol. 31, pp. 110–113; February 14, 1958.
T. C. Chen and A. Papoulis, "Terminal properties of magnetic cores," Proc. IRE, vol. 46, pp. 839–849; May, 1958.
U. F. Gianola, "Switching in rectangular loop ferrites containing air gaps," *J. Appl. Phys.*, vol. 29, pp. 1122–1124; July, 1958.
Proc. 4th Symp. on Magnetism and Magnetic Materials, Philadelphia, Pa.; November 17–20, 1958. *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 1S–322S, April, 1959.

3. Flux Reversal Physics

M. Wolman and H. Kaden, "Eddy current delay of switching phenomena," *Z. Tech. Phys.*, vol. 13, pp. 330–335; July, 1932.
F. J. Beck and M. J. Kelly, "Magnetization in perpendicularly superposed direct and alternating fields," *J. Appl. Phys.*, vol. 19, pp. 551–562; June, 1948.
C. Kittel, "Physical theory of ferromagnetic domains," *Rev. Mod. Phys.*, vol. 21, pp. 541–583; 1949 (lists 67 references).
R. M. Bozorth, "Ferromagnetism," D. Van Nostrand Co., Inc., Princeton, N. J.; 1951.
J. K. Galt, J. Andrus, and H. G. Hooper, "Motion of domain walls in ferrite crystals," *Rev. Mod. Phys.*, vol. 25, pp. 93–97; January, 1953.
J. B. Goodenough, "A theory of domain creation and coercive force in polycrystalline ferromagnetics," *Phys. Rev.*, vol. 95, pp. 917–932; August 15, 1954.
C. Kittel, "Magnetism" in "Modern Physics for the Engineer," L. N. Ridenour, ed., McGraw-Hill Book Co., Inc., New York, N. Y.; 1954.
N. Menyuk and J. B. Goodenough, "Magnetic materials for digital computer components, I. A theory of flux reversal in polycrystalline ferromagnetics," *J. Appl. Phys.*, vol. 26, pp. 8–18; January, 1955.
C. P. Bean and D. S. Rodbell, "Kinetics of magnetization in some square-loop magnetic tapes," *J. Appl. Phys.*, vol. 26, pp. 124–125; January, 1955.
R. L. Conger, "Magnetization reversal in thin films," *Phys. Rev.*, vol. 98, pp. 1752–1754; June 15, 1955.
R. Hertz and H. O. Buelteman, Jr., "The application of perpendicularly superposed magnetic fields," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 74), pp. 655–660; November, 1955. See also, *Elec. Engrg.*, vol. 74, p. 661; August, 1955.
R. W. Roberts and R. L. Van Nice, "Influence of ID-OD ratio on magnetic properties of toroidal cores," *Elec. Engrg.*, vol. 74, pp. 910–914; October, 1955.
D. S. Rodbell and C. P. Bean, "Influence of pulsed magnetic fields on the reversal of magnetization in square-loop metallic tape," *J. Appl. Phys.*, vol. 26, pp. 1318–1323; November, 1955.
C. A. Fowler, Jr. and E. M. Fryer, "Magnetic domains in thin films of nickel-iron," *Phys. Rev.*, vol. 100, p. 746; 1955.
C. B. Wakeman and F. J. Beck, "Superposed magnetic fields in materials with rectangular hysteresis loops," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 75), pp. 562–569; November, 1956.
D. M. Grimes, "Reversible Properties of Ferromagnets," Library of Congress, Washington 25, D. C., PB-126868, 58 pp.; December, 1956.
P. E. Tannenwald, "Advances in solid-state electronics-magnetics," *Proc. Natl. Electronics Conf.*, Chicago, Ill., vol. 12, pp. 741–753; 1956 (contains 26 references).
R. L. Harvey, "Understanding your Magnetic Materials," RCA, Camden, N. J.; April/May, 1957.
R. W. Roberts and C. C. Horstman, "Core tester simplifies ferro-amplifier design," *Electronics*, vol. 30, pp. 150–153; August 1, 1957.
G. Biorci and D. Prescetti, "Observations on the behavior of ferromagnetic materials in toroidal rings," *Energia Elett.*, vol. 34, pp. 795–799; August, 1957.
E. M. Gyorgy, "Rotational model of flux reversal in square-loop ferrites," *J. Appl. Phys.*, vol. 28, pp. 1011–1015; September, 1957.
V. L. Newhouse, "The utilization of domain wall viscosity in data-handling devices," *Proc. WJCC*, pp. 73–80; February, 1957. Also, Proc. IRE, vol. 45, pp. 1484–1494; November, 1957.
V. E. Legg, "Survey of square loop materials," IRE Trans. on Component Parts, vol. CP-4, pp. 106–109; December, 1957.
T. J. Pula, "Volt second transfer efficiency in fast response magnetic amplifiers, Part I—N_2/R and control," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 77), pp. 861–867; (1958) January, 1959. See also, AIEE Publication T-98 (*Proc. Spec. Tech. Conf. on Mag. Ampl.*—1957).
M. A. Rozenblat, "Influence of the dimensions of toroidal cores on their static magnetic properties," *Automatika i Telemekhanika*; August, 1958. (In Russian.)
L. B. Slopian, "On systems of units for electrical and magnetic quantities," *Radiotekhnika*, vol. 13.9, pp. 49–57; September, 1958.
E. M. Gyorgy, "Modified rotational model of flux reversal," *J. Appl. Phys.*, vol. 29, pp. 1709–1712; December, 1958.
C. P. Bean, "Nature of ferromagnetic domain walls," *Bull. Acad. Sci. USSR*, (Phys. Sect.), vol. 21.8, p. 1183; 1958. (In Russian.)
R. E. McMahon, "Impulse switching of ferrites," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 16–17; February, 1959.
H. B. Callen, "Rotational remagnetization of thin films," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 96–97; February, 1959.
W. L. Shevel, Jr., "Millimicrosecond switching properties of ferrite computer elements," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 47S–48S; April, 1959.

B. R. Eichbaum, "An evaluation of a new high speed magnetic ferrite system for use in computer components," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 49S–52S; April, 1959.
W. M. Overn and V. J. Korkowski, "A study of the residual states of ferrite cores in computer memory operation," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 52S–53S; April, 1959.
R. W. McKay, "The reversible component of magnetization," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 56S–57S; April, 1959.

4. Single Aperture Memory Circuits

J. W. Forrester, "Digital information storage in three dimensions using magnetic cores," *J. Appl. Phys.*, vol. 22, pp. 44–48; January, 1951.
M. Kincaid, J. M. Alden, and R. B. Hanna, "Static magnetic memory for low cost computers," *Electronics*, vol. 24, pp. 108–111; January, 1951.
J. D. Goodall and T. Lode, "Multistable magnetic memory techniques," *Radio-Electronic Engrg.*, vol. 17, pp. 3–5; December, 1951.
J. D. Goodall and T. Lode, "Multistable magnetic memory systems," *Proc. Natl. Electronics Conf.*, Chicago, Ill., vol. 7, pp. 378–379; 1951 (abstract).
W. N. Papian, "A coincident current magnetic memory cell for the storage of digital information," Proc. IRE, vol. 40, pp. 475–478; April, 1952.
A. Wang, "Static magnetic memory—its applications to computers and controlling systems," *Proc. ACM*, Pittsburgh, Pa., Meeting, pp. 207–212; May 2–3, 1952.
I. L. Auerbach, "A static magnetic memory system for the ENIAC," *Proc. ACM*, Pittsburgh, Pa., Meeting, pp. 213–222; May 2–3, 1952.
J. A. Rajchman, "Static magnetic matrix memory and switching circuits," *RCA Rev.*, vol. 13, pp. 183–201; June, 1952.
J. P. Eckert, Jr., "A survey of digital computer memory systems," Proc. IRE, vol. 41, pp. 1393–1406; October, 1953.
J. A. Rajchman, "A myriabit magnetic core matrix memory," Proc. IRE, vol. 41, pp. 1407–1421; October, 1953.
I. L. Auerbach, "Fast acting digital memory systems," *Elec. Mfg.*, vol. 52, pp. 100–107; October, 1953, and pp. 136–143; November, 1953.
W. N. Papian, "The MIT magnetic-core memory," *Proc. EJCC*, pp. 37–42; December, 1953.
E. Gelbard and W. Olander, "Ferrite memory devices," *Computers and Automation*, vol. 3.5, pp. 6–7, 13; May, 1954.
R. C. Minnick, "Etched Magnetic Storage System," Harvard Computation Lab., Bell Telephone Labs. Rept. No. 9, "Theory of Switching"; October, 1954.
W. N. Papian, "New ferrite core memory uses pulse transformers," *Electronics*, vol. 28, p. 194; March, 1955.
W. N. Papian, "Applications of ferrites to memory systems," *Proc. Eleventh Annual Meeting, Metal Powder Assoc.*, vol. 11, pp. 183–186; May, 1955.
R. S. DiNolfo, "Multi-Coordinate Selection Systems for Magnetic-Core Storage," M.I.T. Lincoln Lab., Lexington, Mass., Rept. 6R-235; May, 1955.
L. N. Ridenour, "Computer memories," *Scientific American*, vol. 192, pp. 92–100; June, 1955.
C. W. Schultz and G. F. Boesen, "Juke box uses ferrite core memory," *Electronics*, vol. 28, p. 138; October, 1955.
J. Raffel and S. Bradspies, "Experiments on a three-core cell for high-speed memories," 1955 IRE Convention Record, pt. 4, pp. 64–69.
J. D. Childress, "The noise problem in the coincident current memory matrix," *AIEE Conf. on Magnetism and Magnetic Materials* (AIEE Publication T-78), pp. 210–218; 1955.
J. D. Childress, "The Noise Problem in the Coincident Current Memory Matrix," M.I.T. Lincoln Lab., Lexington, Mass., Div. 6, 6M-4153; February 13, 1956.
M. A. Alexander, M. Rosenberg, and R. Stuart-Williams, "Ferrite core memory is fast and reliable," *Electronics*, vol. 29, pp. 158–161; February, 1956.
E. A. Guditz and L. B. Smith, "Vacuum and vibration speed assembly of core memory planes," *Electronics*, vol. 29, pp. 214–228; February, 1956.
F. Fowler, Jr., "The computer's memory," *Control Engrg.*, vol. 3, pp. 93–101; May, 1956.
W. J. Bartik and T. H. Bonn, "A small coincident-current magnetic memory," IRE Trans. on Electronic Computers, vol. EC-5, pp. 73–78; June, 1956.
"Magnetic core components for digital control applications," *Automatic Control*, pp. 34–40; July, 1956.
R. E. McMahon and F. L. McNamara, "Transistor Core Memory," M.I.T. Lincoln Lab., Lexington, Mass., Tech. Rept. No. 133; November 13, 1956.
L. J. Andrews, "A technique for using memory cores as logical elements," *Proc. EJCC*, pp. 39–46; December, 1956.
J. L. Mitchell and K. H. Olsen, "TX-O, a transistor computer with a 256×256 memory," *Proc. EJCC*, pp. 93–101; December, 1956.
A. V. Pohm and S. M. Rubens, "A compact coincident current memory," *Proc. EJCC*, pp. 120–123; December, 1956.
R. E. Merwin, "The IBM 705 EDPM memory system," IRE Trans. on Electronic Computers, vol. EC-5, pp. 219–223; December, 1956.
A. E. Joel, "Electronics in telephone switching systems," (*Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 75), pp. 701–710; (1956), p. 707; January, 1957.
R. O. Endres, "High-speed memory for symposium computer," *Proc. Symp. on Computer Fundamentals*, AIEE-IRE, Philadelphia, Pa.; January/March, 1957.
G. E. Valenty, "A medium-speed magnetic core memory," *Proc. WJCC*, pp. 55–67; February, 1957.
V. L. Newhouse, "The utilization of domain wall viscosity in data-handling devices," *Proc. WJCC*, pp. 73–80; February, 1957. Also, Proc. IRE, vol. 45, pp. 1484–1494; November, 1957.
A. Katz, A. G. Jones, and G. Rezek, "Accuracy control system for magnetic core memories," *Proc. WJCC*, pp. 105–109; February, 1957.
R. L. Best, "Memory units in the Lincoln TX-2," *Proc. WJCC*, pp. 160–167; February, 1957.
J. A. Rajchman, "Ferrite aperture plate for random access memory," Proc. IRE, vol. 45, pp. 325–334; March, 1957.
E. D. Foss and R. S. Partridge, "A 32,000-word magnetic core memory," *IBM J. Res. and Dev.*, vol. 1, pp. 102–104; April, 1957.
E. A. Guditz, "Three-dimensional printed wiring," *Electronics*, vol. 30, pp. 160–163; June, 1957.
F. L. McNamara, "The noise problem in a coincident current core memory," IRE Trans. on Instrumentation, vol. I-6, pp. 153–156; June, 1957.
R. E. McMahon, "Transistorized core memory," IRE Trans. on Instrumentation, vol. I-6, pp. 157–160; June, 1957.
A. A. Ashley, "A five-microsecond memory for (the) UDOFT computer," 1957 WESCON Convention Record, pt. 4, pp. 262–266.
J. A. Rajchman, "Principles of Transfluxor and Core Circuits," RCA Labs. Div., Princeton, N. J., RCA Industry Service Lab. Bull. RB-120; September 10, 1957.
W. N. Papian, "High speed computer stores 2.5 megabits," *Electronics*, vol. 30, pp. 162–167; October 1, 1957.
L. S. Michels, "A unique approach to computer versatility," *Tele-Tech. & Electronic Ind.*, vol. 16, pp. 75–77, 142; October, 1957.
R. K. Richards, "Digital Computer Components and Circuits," D. Van Nostrand Co., Inc., Princeton, N. J.; 1957.
W. Renwick, "A magnetic core matrix store with direct selection using a magnetic core switch matrix," *Proc. IEE*, vol. 104, pt. B, suppl. no. 7, pp. 436–444, 481–482; 1957.
C. S. Warren, W. G. Rumble, and W. A. Helbig, "Transistorized memory monitors earth satellite," *Electronics*, vol. 31, pp. 66–70; January 17, 1958.
H. Small, "The magnetic core as a storage device," *Philco Tech. Rept. Bull.*, vol. 8.2, pp. 3–6; March/April, 1958.
H. Small, "Digital computer memory requirements," *Philco Tech. Rept. Bull.*, vol. 8.3, pp. 12–19; May/June, 1958.
J. A. Rajchman, "Magnetic switching," *Proc. WJCC*; May, 1958.
B. Goda, S. Markowitz, and E. Jacobi, "A line printer memory utilizing magnetic cores," *Computers and Automation*, vol. 7, pp. 12–14; August, 1958.
D. H. Looney, "The ferrite bead—a new memory device," *Proc. AIEE Conf. on Magnetism and Magnetic Materials*, Philadelphia, Pa.; November, 1958.
C. C. Gotlieb and J. N. P. Hume, "High Speed Data Processing," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 39–41; 1958.
R. Stuart-Williams and M. A. Alexander, "Recent Advances in Coincident Current Magnetic Memory Techniques," Intl. Telemeter Corp., Los Angeles, Calif.; 1958.

5. Memory Access Circuits

J. A. Rajchman, "Static magnetic matrix memory and switching circuits," *RCA Rev.*, vol. 13, pp. 183–201; June, 1952.

K. H. Olsen, "A Magnetic Matrix Switch and Its Incorporation into a Coincident Current Memory," Digital Computer Lab., M.I.T., Cambridge, Mass., Rept. R-211; June, 1952.
M. K. Haynes, "Multidimensional magnetic memory selection systems," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-1, pp. 25–29; December, 1952.
R. L. Ashenhurst and R. C. Minnick, "Multiple Coincidence Magnetic Storage Systems," Harvard Computation Lab., Bell Telephone Labs. Rept. No. 4, "Theory of Switching"; July, 1953.
J. A. Rajchman, "A myriabit magnetic core matrix memory," PROC. IRE, vol. 41, pp. 1407–1421; October, 1953.
H. Billing, "Schnell aufrufbare magnetostatische Speicher fuer elektronische Rechenmaschinen," *Naturwissenschaften*, vol. 40.2, p. 49; 1953.
R. C. Minnick, "Output Wire Considerations for Pierced Multiple-Coincidence Magnetic Storage Systems," Harvard Computation Lab., Bell Telephone Labs. Rept. No. 7, "Theory of Switching"; May, 1954.
J. I. Raffel, "Switch for Register Selection in a Magnetic Core Memory," M.I.T. Lincoln Lab., Lexington, Mass., Rept. R-234; July 27, 1954.
R. S. DiNolfo, "Multi-coordinate selection systems for magnetic core storage," Master's thesis, M.I.T., Cambridge, Mass.; September, 1954.
R. C. Minnick, "Etched Magnetic Storage System," Harvard Computation Lab., Bell Telephone Labs. Rept. No. 9, "Theory of Switching"; October, 1954.
J. M. Weir, "A high-speed permanent storage device," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-4, pp. 16–20; March, 1955.
R. C. Minnick and R. L. Ashenhurst, "Multiple coincidence magnetic storage systems," *J. Appl. Phys.*, vol. 26, pp. 575–579; May, 1955.
R. S. DiNolfo, "Multi-Coordinate Selection Systems for Magnetic Core Storage," M.I.T. Lincoln Lab., Lexington, Mass., Rept. R-235; June, 1955.
E. J. Otis, "Series Tube Core Memory," USAF AFCRC-TB-55-116, U. S. Dept. of Commerce Publication PB-121558; September, 1955.
E. A. Guditz and L. B. Smith, "An improved technique for the assembly of core memory planes," *Electronics*, vol. 29.2, pp. 214–228; February, 1956.
N. M. Blachman, "On the wiring of two-dimensional multiple-coincidence magnetic memories," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-5, pp. 19–21; March, 1956.
F. S. Goulding, "Some transistor circuits used in a magnetic core type kicksorter," 1956 IRE CONVENTION RECORD, pt. 9, pp. 76–81.
A. A. Robinson, V. L. Newhouse, M. J. Friedman, D. G. Bindon, and I. P. V. Carter, "A digital store using a magnetic core matrix," *Proc. IEE*, vol. 103, pt. B, supplement 2, pp. 295–301; April, 1956.
H. Gillert, "Bemessung der Bauelemente eines Schnellspeichers mit Ferrit-Ringkernen (Design of the components of a fast acting store using ferrite ring cores)," *Nachrichtentech Z.*, vol. 9, pp. 250–252; June, 1956.
W. J. Bartik and T. H. Bonn, "A small coincident current magnetic memory," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-5, pp. 73–78; June, 1956.
S. Bradspeis, "A Transistorized Sensing Amplifiers for the 256×256 Core Memory," M.I.T. Lincoln Lab., Lexington, Mass., Internal Rept. 16M-4785; November, 1956.
R. E. McMahon and F. L. McNamara, "Transistor Core Memory," M.I.T. Lincoln Lab., Lexington, Mass., Tech. Rept. No. 133; November 13, 1956.
R. E. Merwin, "The IBM 705 EDPM memory system," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-5, pp. 219–223; December, 1956.
C. S. Warren, W. G. Rumble, and W. A. Helbig, "A transistor operated magnetic core memory," *Proc. Natl. Electronics Conf.*, Chicago, Ill., vol. 12, pp. 289–299; 1956.
W. E. Bradley, "Transistorized ferrite core buffer memory," *Proc. Natl. Electronics Conf.*, Chicago, Ill., vol. 12, pp. 308–311; 1956.
Bell Telephone Labs., "Transistorized magnetic core memory," *Computers and Automation*, vol. 6, pp. 26–27; January, 1957.
R. O. Endres, "High-speed memory for symposium computer," *Proc. Symp. on Computer Fundamentals*, Paper No. 5, AIEE/IRE, Philadelphia, Pa.; January/March, 1957.
J. A. Rajchman and H. D. Crane, "Current steering in magnetic circuits," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-6, pp. 21–30; March, 1957.
W. A. Helbig and W. G. Rumble, "How to select transistor switches for magnetic memories," *Electronic Equipment*, pp. 22–25; March, 1957. (Based on a 1956 NEC paper; see above.)
E. L. Younker, "A transistor-driven magnetic core memory," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-6, pp. 14–20; March, 1957.
J. D. Lawrence and T. H. Bonn, "A magnetic pulse-current regulator," 1957 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 102–105.
E. D. Foss and R. S. Partridge, "A 32,000-word magnetic core memory," *IBM J. Res. and Dev.*, vol. 1, pp. 102–109; April, 1957.
W. W. Peterson, "Addressing for random-access storage," *IBM J. Res. and Dev.*, vol. 1, pp. 116–129; April, 1957.
R. C. Minnick, "Simultaneous access matrix switching systems," paper presented at the Internatl. Symp. on Switching Theory, Harvard University, Cambridge, Mass.; April 2–5, 1957.
R. E. McMahon, "Transistorized core memory," IRE TRANS. ON INSTRUMENTATION, vol. I-6, pp. 157–160; June, 1957.
A. M. Angel, "A very high-speed punched paper tape reader," 1957 WESCON CONVENTION RECORD, pt. 4, pp. 218–226.
R. E. McMahon, "Survey of Memory Techniques Using Transistors," M.I.T. Lincoln Lab., Lexington, Mass., Rept. 2G-24-82; October 1, 1957.
R. K. Richards, "Digital Computer Components and Circuits," D. Van Nostrand Co., Inc., Princeton, N. J.; 1957.
W. Renwick, "A magnetic core matrix store with direct selection using a magnetic core switch matrix," *Proc. IEE*, vol. 104, pt. B, suppl. no. 7, pp. 436–444, 481–482; 1957.
C. S. Warren, W. G. Rumble, and W. A. Helbig, "Transistorized memory monitors earth satellite," *Electronics*, vol. 31.3, pp. 66–70; January 17, 1958.
R. E. McMahon, "Linear selection core-memory techniques using transistors," paper presented at the 1958 Transistor and Solid-State Circuits Conf., Philadelphia, Pa.; February 20–21, 1958.
C. Cottini, E. Gatti, and G. Giannelli, "Transistor current generator for fast ferrite memory matrix," *Nuclear Instr.*, vol. 2, p. 164; February, 1958.
A. L. Lane and A. Turczyn, "A high-speed-*n*-pole, *n*-position magnetic core matrix switch," 1958 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 246–253.
W. J. Haneman, J. Lehmann, and C. S. Warren, "Apertured plate memory: operation and analysis," 1958 IRE NATIONAL CONVENTION RECORD, pt. 4, p. 254 (abstract).
G. Constantine, Jr., "A load-sharing matrix switch," *IBM J. Res. and Dev.*; July, 1958.
R. L. Koppel, "A transistorized driven magnetic core memory using non-coincident current techniques," *Proc. 1958 Natl. Symp. on Telemetering*, paper 1.3.
J. A. Rajchman, G. R. Briggs, and A. W. Lo, "Transfluxor controlled electroluminescent display panels," PROC. IRE, vol. 46, pp. 1808–1824; November, 1958.
G. F. Abbott, Jr., "High-current, high-speed, non-saturating transistor drivers," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa.; February, 1959.
R. E. McMahon, "Impulse switching of ferrites," *Digest of Tech. Papers, Proc. 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 16–17; February, 1959.

6. MULTI-APERTURE MAGNETIC DEVICES

J. A. Rajchman and A. W. Lo, "The transfluxor—a magnetic gate with stored variable setting," *RCA Rev.*, vol. 16, pp. 303–311; June, 1955.
R. L. Snyder, "Magnistor circuits," *Elect. Design*, vol. 3; August, 1955.
J. A. Rajchman and A. W. Lo, "The transfluxor," *Proc. WJCC*, pp. 109–118; February, 1956. Also, PROC. IRE, vol. 44, pp. 321–332; March, 1956.
L. P. Hunter and E. W. Bauer, "High speed concident flux storage principles," *J. Appl. Phys.*, vol. 27, pp. 1257–1261; November, 1956.
W. W. Lawrence, Jr., "Recent development in very-high speed magnetic storage techniques," *Proc. EJCC*, pp. 101–103; December, 1956.
A. Darré, "Abfragen Magnetischer Speicher ohne Informationsverlust," *Frequenz*, vol. 11, pp. 19–27; January, 1957, and pp. 38–42; February, 1957.
J. A. Rajchman and H. D. Crane, "Current steering in magnetic circuits," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-6, pp. 21–30; March, 1957.

H. W. Abbott and J. J. Suran, "Temperature characteristics of the transfluxor," IRE TRANS. ON ELECTRON DEVICES, vol. ED-4, pp. 113–119; April, 1957.
L. S. Onyshkevych, "Analysis of circuits with multiple-hole magnetic circuits," Master's thesis, M.I.T., Cambridge, Mass.; May, 1957.
R. B. Goldner, "Application of transfluxors to selective calling," Master's thesis, M.I.T., Cambridge, Mass.; Spring, 1957.
H. W. Abbott and J. J. Suran, "Multihole ferrite core configurations and applications," PROC. IRE, vol. 45, pp. 1081–1093; August, 1957.
J. A. Rajchman, "A survey of magnetic and other solid-state devices for the manipulation of information," IRE TRANS. ON CIRCUIT THEORY, vol. CT-4, pp. 210–225; September, 1957.
J. A. Rajchman, "Principles of Transfluxor and Core Circuits," RCA Labs. Div., Princeton, N. J., RCA Industry Service Lab. Bull. RB-120; September 10, 1957.
E. W. Bauer, "Nondestructive Sensing of Magnetic Cores," U. S. Patent No. 2,814,794; November 26, 1957 (assigned to IBM).
S. A. Abbas and D. L. Critchlow, "Calculation of flux patterns in ferrite multipath structures," 1958 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 263–267.
N. F. Lockhart, "Logic by ordered flux changes in multipath ferrite cores," 1958 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 268–278.
J. A. Rajchman, "Magnetic switching," *Proc. WJCC*; May, 1958.
H. D. Crane, "A high-speed logic system using magnetic elements and connecting wire only," PROC. IRE, vol. 47, pp. 63–73; January, 1959 (see also *Proc. Tech. Conf. on Nonlinear Magnetics and Magnetic Amplifiers*, Los Angeles, Calif., pp. 465–482; August 6–8, 1958).
"New device may replace magnetic core," *Electronic News*, pp. 1, 18; August 18, 1958.
"New uses for versatile magnetor," *Automatic Control*, p. 49; November, 1958.
J. A. Rajchman, G. R. Briggs, and A. W. Lo, "Transfluxor controlled electroluminescent display panels," PROC. IRE, vol. 46, pp. 1808–1824; November, 1958.
J. A. Fusca, "Ferrite cores perform logic functions," *Aviation Week*, vol. 69.19, pp. 83–92; November 10, 1958.
"LADDIC," *Fortune*, vol. 57, no. 5, p. 156; November, 1958.
"Magnetic device for logic circuitry," *J. Appl. Phys.*, vol. 29, cover and p. 1776; December, 1958.
N. S. Prywes, "Diodeless magnetic shift registers utilizing transfluxors," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 316–324; December, 1958.
U. F. Gianola and T. H. Crowley, "The Laddic, a magnetic device for performing logic," *Bell Sys. Tech. J.*, vol. 38, pp. 45–72; January, 1959.
R. Henkel, "Multi-aperture magnetic memory unit detailed," *Electronic News*, vol. 4.132, pp. 1 and 18; March 9, 1959.
D. G. Hammel, W. L. Morgan, and R. D. Sidnam, "A multiload transfluxor memory," *Proc. WJCC*; March, 1959.
H. D. Crane and D. R. Bennion, "Design and analysis of MAD transfer circuitry," *Proc. WJCC*; March, 1959.
D. H. Looney, "Recent advances in magnetic devices for computers," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 38S–42S; April, 1959.
D. A. Meier, "Millimicrosecond magnetic switching and storage element," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 45S–46S; April, 1959.
J. A. Baldwin, Jr. and J. L. Rogers, "Inhibited flux—a new mode of operation of the three-hole memory core," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 58S–59S; April, 1959.

7. THIN FILM MEMORY ELEMENTS

(See Also Section 1c, Thin Film Materials)

R. L. Conger, "Magnetization reversal in thin films," *Phys. Rev.*, vol. 98, pp. 1752–1754; June 15, 1955.
C. A. Fowler, Jr. and E. M. Fryer, "Magnetic domains in thin films of nickel-iron," *Phys. Rev.*, vol. 100, p. 746; 1955.
D. A. Buck, "The cryotron—a superconductive computer component," PROC. IRE, vol. 44, pp. 482–493; April, 1956.
A. E. Slade and H. O. McMahon, "Cryotron catalog memory system," *Proc. EJCC*, pp. 115–119; December, 1956.
A. V. Pohm and S. M. Rubens, "A compact coincident current memory," *Proc. EJCC*, pp. 120–123; December, 1956.
J. W. Crowe, "Trapped flux superconducting memory," *IBM J. Res. and Dev.*, vol. 1, pp. 294–303; October, 1957.
W. Arrott, "Design advances in magnetism research," *Elec. Mfg.*, pp. 95–97; January, 1958.
J. W. Bremer, "Cryogenic devices in logical circuitry and storage," *Elec. Mfg.*, pp. 78–83; February, 1958.
E. E. Bittmann, "Thin film memories," presented at *1959 Solid-State Circuits Conf.*, Philadelphia, Pa.; February, 1959. IRE TRANS. ON ELECTRONIC COMPUTERS, this issue.
C. J. Kraus, "Pros and cons on a superconducting memory," presented at *1959 Solid-State Circuits Conf.*, Philadelphia, Pa., February, 1959; to be published in IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-8; September, 1959.
L. J. Oakland and T. D. Rossing, "Coincident-current nondestructive readout from thin magnetic films," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 54S–55S; April, 1959.
J. I. Raffel, "Operating characteristics of a thin film memory," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 60S–61S; April, 1959.

8. "TWISTORS"

A. H. Bobeck, "A new storage element suitable for large-sized memory arrays—the twistor," *Bell Sys. Tech. J.*, vol. 36, pp. 1319–1340; November, 1957.
W. Arrott, "Design advances in magnetism research," *Elec. Mfg.*, pp. 95–97; January, 1958.
U. F. Gianola, "Nondestructive memory employing a domain orientated steel wire," *J. Appl. Phys.*, vol. 29, pp. 849–853; May, 1958.
A. H. Bobeck and R. S. Title, "The twistor—its use in a memory array," *Proc. Conf. on Nonlinear Magnetics and Magnetic Amplifiers*, Los Angeles, Calif.; August 6–8, 1958.
S. J. Schwartz and J. S. Sallo, "Electrodeposited twistor and bit wire components," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 94–95; February, 1959.
G. F. Abbott, Jr., High-current, high-speed, nonsaturating transistor drivers," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 64–65; February, 1959.
K. Preston, Jr. and Q. W. Simkins, "Twistor buffer store," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 14–15; February, 1959.
A. H. Bobeck and R. F. Fischer, "A reversible, diodeless, twistor shift register," *J. Appl. Phys.*, suppl. to vol. 30.4, pp. 43S–44S; April, 1959.

9. FERRITE APERTURE PLATES

J. A. Rajchman, "Ferrite aperture plate for random access memory," *Proc. EJCC*, pp. 107–115; December, 1956.
R. L. Meinken, "A memory array in a sheet of ferrite," *Proc. Conf. on Magnetism and Magnetic Materials*, p. 674; 1956 (AIEE Publication T-91).
J. A. Rajchman, "Ferrite aperture plate for random access memory," PROC. IRE, vol. 45, pp. 325–334; March, 1957.
W. J. Haneman, J. Lehmann, and C. S. Warren, "Apertured plate memory: operation and analysis," 1958 IRE NATIONAL CONVENTION RECORD, pt. 4, p. 254 (abstract).
M. M. Kaufman and V. L. Newhouse, "Operating range of a memory using two ferrite plate apertures per bit," *J. Appl. Phys.*, p. 487; March, 1958.
V. L. Newhouse, *et al.*, "End fired memory uses ferrite plates," *Electronics*, 4 pp.; October 10, 1958.
C. S. Warren, "Ferrite aperture plate memories," *Digest of Tech. Papers, 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 18–19; February, 1959.

10. DOMAIN WALL VISCOSITY MEMORIES

V. L. Newhouse, "The utilization of domain wall viscosity in data-handling devices," *Proc. WJCC*, pp. 73–80; February, 1957. Also, PROC. IRE, vol. 45, pp. 1484–1494; November, 1957.
C. P. Bean, "Nature of ferromagnetic domain walls," *Bull. Acad. Sci. USSR* (Physics Sect.), vol. 21.8, p. 1183; 1958. (In Russian.)

11. RF METHODS (PARAMETRON AND FERRISTOR)

C. Isborn, "Ferroresonant flip-flops," *Electronics*, vol. 25, p. 121; April, 1952.
B. Woodrow, "A radio-frequency nondestructive read-out of magnetic core memories," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-3, pp. 12–15; December, 1954.
E. Swarthe and C. Isborn, "Counting labels with ferroresonant rings," *Electronics*, vol. 28, pp. 128–129; October, 1955.
E. Goto, "On the application of parametrically excited nonlinear resonators," *J. Inst. Elec. Comm. Engrs. of Japan*, vol. 38, pp. 770 ff.; October, 1955. (In Japanese.)

C. B. Newport and D. A. Bell, "Ferroresonant circuits for digital computers," *J. Brit. IRE*, vol. 17, pp. 619–630; November, 1957.
A. H. Argabrite, "Subminiature magnetic amplifiers," *Radio and TV News*, vol. 58.6, pp. 70–71, 175; December, 1957.
Staff of the (Harvard) Computation Lab., "Description of the Mark IV Calculator," *Annals of the Computation Lab.*, vol. 28.
S. Muroga, "Elementary principle of parametron and its application to digital computers," *Datamation*, vol. 4.5, pp. 31–34; September/October, 1958.
G. E. Kelly, Jr., "The ferroresonant circuit," *Commun. & Electronics*, no. 40 (*Trans. AIEE*, pt. I, vol. 77), pp. 843–848; (1958) January, 1959.
R. H. Dennard, "Behavior of the ferroresonant series circuit containing a square-loop reactor," *Commun. & Electronics*, no. 40 (*Trans. AIEE*, pt. I, vol. 77), pp. 903–911; (1958) January, 1959.
E. Goto, "The parametron, a digital computing element which utilizes parametric oscillation," to be published in PROC. IRE; August, 1959.

12. Cross-Field Effects

D. A. Buck and W. I. Frank, "Nondestructive sensing of magnetic cores," *Commun. & Electronics*, no. 10 (*Trans. AIEE*, pt. I, vol. 72), pp. 822–830; January, 1954.
R. H. Fuller, "A Static-Magnetic Data-Storage Unit," M.I.T. Servo Lab., Rept. 9; May 1, 1954.
A. Papoulis, "The nondestructive read-out of magnetic cores," PROC. IRE, vol. 42, pp. 1283–1288; August, 1954.
R. Thorensen and R. W. Arsenault, "A new nondestructive read for magnetic cores," *Proc. WJCC*, pp. 111–115; March, 1955.
W. Matthews, "Telemetering in earth satellites," *Elec. Engrg.*, vol. 76, pp. 976–981; November, 1957.
D. H. Schaefer, "Magnetic core event counter for earth satellite memory," *Elec. Engrg.*, vol. 77, pp. 52–56; January, 1958.
W. Arrott, "Design advances from magnetism research," *Elec. Mfg.*, p. 95; January, 1958.
C. B. House and R. L. Van Allen, "Commutation and Nondestructive Readout of Magnetic Memory Cores in Earth Satellite," N.R.L., Washington, D. C., Natl. Telemetering Conf. Rept.; 1958.
T. D. Rossing and S. M. Rubens, "Effects of a transverse field on switching rates of magnetic memory cores," *J. Appl. Phys.*, vol. 29, pp. 1245–1247; August, 1958.

13. Magnetic Amplifiers (Ramey-Type Circuits)

(See also Section 18, Other Bibliographies)

R. A. Ramey, "The Single Core Magnetic Amplifier as a Computer Element," N.R.L. Rept. 4030; August 8, 1952.
R. A. Ramey, "The single core magnetic amplifier as a computer element," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 71), pp. 342–346; (1952) January, 1953.
C. W. Lufey, "A survey of magnetic amplifiers," PROC. IRE, vol. 43, pp. 404–413; April, 1955.
A. J. Lincoln, "Ferromagnetic Core Logical Circuitry and its Application to Digital Computers," Office of Tech. Services U. S. Dept. of Commerce, Washington, D. C., PB111954; August, 1955.
T. H. Bonn, "The megacycle ferractor," *Proc. WJCC*; February, 1956.
R. D. Torrey and A. L. Krell, "Series Pulse Magnetic Amplifiers," Sci. Rept. AFC-RC-TN-15-186, AD110190; February 1, 1957.
R. E. Morgan, "Reactions of bistable B-H loops on magnetic amplifiers," *Commun. & Electronics*, no. 39 (*Trans. AIEE*, pt. I, vol. 77), pp. 586–590; November, 1958.
AIEE Committee Report, "Magnetic amplifier bibliography, 1951–1956," *Commun. & Electronics*, no. 39 (*Trans. AIEE*, pt. I, vol. 77), pp. 613–627; November, 1958.
J. R. Walker, "Proposed AIEE terms and definitions for magnetic amplifiers," *Electronic Design News*; January, 1959.

14. Other Memory Systems

J. Raffel and S. Bradspies, "Experiments on a three-core cell for high-speed memories," 1955 IRE CONVENTION RECORD, pt. 4, pp. 64–69.
R. A. Tracy, "Megabit memory," *Proc. EJCC*, pp. 104–106; December, 1956.
R. L. Koppel, "A transistor driven magnetic core memory using non-coincident current techniques," *Proc. 1958 Natl. Symp. on Telemetering*, paper 1.3; September, 1958.

15. Magnetic Logic Circuits

A. Wang and W. D. Woo, "Static magnetic storage and delay line," *J. Appl. Phys.*, vol. 21, pp. 49–54; January, 1950.
A. Wang, "Magnetic triggers," PROC. IRE, vol. 38, pp. 626–629; June, 1950.
M. K. Haynes, "Magnetic cores as elements of digital computing systems," Ph.D. dissertation, University of Illinois, Urbana, August 28, 1950.
A. Wang, "Magnetic delay line storage," PROC. IRE, vol. 39, pp. 401–407; April, 1951.
J. G. Miles, "Bibliography of magnetic amplifier devices and the saturable reactor art," *Applications & Industry* (*Trans. AIEE*, pt. II, vol. 70), pp. 2104–2123; 1951.
N. B. Saunders, "Magnetic binaries in the logical design of information handling machines," *Proc. ACM*, Pittsburgh, Pa., Meeting, May 2–3, 1952, pp. 223–229; May, 1952.
G. R. Briggs, "A Magnetic Core Gate and its Application in a Stepping Register," M.I.T. Digital Computer Lab. Engrg. Note E-475; October, 1952.
R. A. Ramey, "The single core magnetic amplifier as a computer element," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 71), pp. 442–446; (1952) January, 1953.
D. R. Brown, "Magnetic materials for high speed pulse circuits," paper presented at the AIEE Winter General Meeting, New York, N. Y.; January, 1953. Also, M.I.T. Digital Computer Lab. Engrg. Note E-530; February 27, 1953.
R. D. Kodis, S. Ruhman, and W. D. Woo, "Magnetic shift register using one core per bit," 1953 IRE CONVENTION RECORD, pt. 7, pp. 38–42.
D. R. Brown and E. Albers-Schoenberg, "Ferrites speed digital computers," *Electronics*, vol. 26, pp. 146–149; April, 1953.
E. A. Sands, "An analysis of magnetic shift register operations, PROC. IRE, vol. 41, pp. 993–998; August, 1953.
H. K. Rising, "Magnetic-core pulse amplifiers for digital computer applications," Master's thesis, M.I.T., Cambridge, Mass.; 1953.
S. Guterman and R. D. Kodis, "Magnetic core ring counter," *Proc. Natl. Electronics Conf.*, Chicago, Ill., vol. 9, pp. 665–669; 1953.
J. D. Goodell, "Testing magnetic decision elements," *Electronics*, vol. 27, pp. 200–203; January, 1954.
D. H. Jacobs, "Design features of the Jaincomp-C and Jaincomp-D electronic digital computers," 1954 IRE CONVENTION RECORD, pt. 4, pp. 98–104.
S. Guterman and R. D. Kodis, "Circuits to perform logical and control functions with magnetic cores," 1954 IRE CONVENTION RECORD, pt. 4, pp. 124–132.
R. C. Minnick, "Magnetic switching circuits," *J. Appl. Phys.*, vol. 25, pp. 479–485; April, 1954.
V. L. Newhouse, "Magnetic and ferroelectric computing components," *Electronic Engrg.*, pp. 192–199; May, 1954.
R. D. Kodis, "Application and performance of magnetic core circuits in computing systems," *Proc. EJCC*, pp. 30–34; December, 1954.
S. Guterman, R. D. Kodis, and S. Ruhman, "Logical and control functions performed with magnetic cores," PROC. IRE, vol. 43, pp. 291–298; March, 1955.
W. Miehle, J. Paivinen, J. Wylen, and D. Loev, "Bimag circuits for digital data processing," 1955 IRE CONVENTION RECORD, pt. 4, pp. 70–83.
S. Guterman and W. M. Carey, "A transistor-magnetic core circuit; a new device applied to digital computing techniques," 1955 IRE CONVENTION RECORD, pt. 4, pp. 84–94.
C. W. Lufey, "A survey of magnetic amplifiers," PROC. IRE, vol. 43, pp. 404–413; April, 1955.
R. D. Kodis, "Using magnetic cores in computers," *Electronic Design*, pp. 22–25; April, 1955.
C. A. Lay, Jr., "A Single Line Magnetic Shift Register Using Diode Amplifiers," M.I.T. Servo Lab. Engrg. Rept. No. 12, D.I.C. 6873 (AF 33 (038)-24007); May, 1955.
M. Karnaugh, "Pulse switching circuits using magnetic cores," PROC. IRE, vol. 43, pp. 570–584; May, 1955 (contains 31 references).
A. J. Lincoln, "Ferromagnetic Core Logical Circuitry and its Application to Digital Computers," Office of Tech. Services, Dept. of Commerce, Washington, D. C., PB 111954; August, 1955.
I. L. Auerbach and S. B. Disson, "Magnetic elements in arithmetic and control circuits," *Elec. Engrg.*, vol. 74, pp. 766–770; September, 1955.
E. Goto, "On the application of parametrically excited non-linear resonators," *J. Inst. Elec. Comm. Engrs. of Japan*, vol. 38, p. 770ff.; October, 1955. (In Japanese.)

E. Swarthe and C. Isborn, "Counting labels with ferroresonant rings," *Electronics*, vol. 28, pp. 128–129; October, 1955.
L. A. Finzi and G. C. Feth, "Magnetic amplifiers in bistable operation," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 74), pp. 592–598; November, 1955.
D. Loev, W. Miehle, J. Paivinen, and J. Wylen, "Magnetic core circuits for digital data-processing systems," PROC. IRE, vol. 44, pp. 154–162; February, 1956.
T. H. Bonn, "The megacycle ferractor," *Proc. WJCC*, pp. 118–119; February, 1956.
N. R. Scott, "Practical circuits for gating in digital computers," *Control Engrg.*, vol. 3, pp. 93–98; February, 1956.
D. A. Weinstein, "The application of magnetic techniques to a reliable 40 kc EPUT meter design," 1956 IRE CONVENTION RECORD, pt. 5, pp. 25–31.
F. S. Goulding, "Some transistor circuits used in a magnetic core type kicksorter," 1956 IRE CONVENTION RECORD, pt. 9, pp. 76–81.
D. H. Schaefer and R. L. Van Allen, "Transcendental function analog computation with magnetic cores," *Commun. & Electronics*, no. 24 (*Trans. AIEE*, pt. I, vol. 75), pp. 160–165; May, 1956.
D. Katz, "A magnetic amplifier switching matrix," *Commun. & Electronics*, no. 24 (*Trans. AIEE*, pt. I, vol. 75), pp. 236–241; May, 1956.
"USAF computer uses magnetic amplifiers," *Electronics*, vol. 29, pp. 138–139; July, 1956.
"Magnetic core components for digital control applications," *Automatic Control*, pp. 34–40; July, 1956.
R. C. Kelner and M. H. Glauberman, "Magnetic shift register correlator," *Electronics*, vol. 29, pp. 172–175; August, 1956.
W. H. Bridge, "Magnetic cores as generalized nonlinear components," 1956 WESCON CONVENTION RECORD. Reprints available from Airtronixs, Inc., Washington, D. C.
V. L. Newhouse and N. S. Prywes, "High-speed shift registers using one core per bit," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-5, pp. 114–120; September, 1956.
R. P. Talambiaras, "Flip-Flop Circuits," U. S. Patent No. 2,749,451 (see also *Elec. Design News*, p. 58; October, 1956), granted June 5, 1956.
G. L. Keister, "Transistor magnetic analog multiplier," *Electronics*, vol. 29, pp. 160–163; October, 1956.
D. Ellis and L. Andrews, "Inverted Core Logic and the Synthesis of Boolean Functions Therefrom," Research Rept., NCR-ED Bull. No. 11, National Cash Register Co., Hawthorne, Calif.; October, 1956.
F. Pawlek and K. Reichel, "Progress in magnetic materials and their application," *EEG Mitt.* (*Berlin*), vol. 46, pp. 337–348; November/December, 1956.
L. J. Andrews, "A technique for using memory cores as logical elements," *Proc. EJCC*, pp. 39–46; December, 1956.
D. A. Buck, "A magnetically controlled gating element," *Proc. EJCC*, pp. 47–50; December, 1956. (AIEE Publication T-92.)
R. D. Torrey and T. H. Bonn, "A 2.5 mc ferractor accumulator," *Proc. EJCC*, pp. 50–53; December, 1956.
E. W. Hogue, "A saturable-transformer digital amplifier with diode switching," *Proc. EJCC*, pp. 58–64; December, 1956. (AIEE Publication T-92.)
J. Keilson and G. Smoliar, "Synchronization of a magnetic computer," *Proc. EJCC*, pp. 90–93; December, 1956.
O. J. Van Sant, Jr., "Zero signal reliability in magnetic shift registers," AIEE Pacific General Meeting Conf., Paper CP-56-694; 1956.
J. Millman and H. Taub, "Pulse and Digital Circuits," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 425–427; 1956.
V. L. Newhouse, "The utilization of domain wall viscosity in data handling devices," *Proc. WJCC*, pp. 73–80; February, 1957. Also, PROC. IRE, vol. 45, pp. 1484–1494; November, 1957.
L. A. Russell, "Diodeless magnetic core logical circuits," 1957 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 106–114.
E. A. Sands, "Magnetically controlled counters," 1957 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 173–174.
J. A. Rajchman and H. D. Crane, "Current steering in magnetic circuits," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-6, pp. 21–30; March, 1957.
J. G. Santesmases, "Switching research in Spain," paper presented at the Internatl. Symp. on Switching Theory, Harvard University, Cambridge, Mass.; April 2–5, 1957.
W. D. Woo, "Magnetic core logical circuits," presented at the Internatl. Symp. on Switching Theory, Harvard University, Cambridge, Mass.; April 2–5, 1957.
T. E. Bray and B. Silverman, "Shaping magnetic and dielectric hysteresis loops," paper presented at AIEE Conf. on Solid-State Dielectrics and Magnetic Devices, Washington, D. C.; April 22–23, 1957.
"How to use magnetic shift registers," *Electronic Equipment*, pp. 6–9; May, 1957.
Newell, "The Design of Logical Circuits Using Transistors and Square-Loop Ferrite Cores," Mullard Ltd. (England); July 1, 1957.
T. H. Bonn, "Magnetic computer has high speed," *Electronics*, vol. 30, pp. 156–160; August 1, 1957.
K. Ganzhorn, "Basic magnetic logic circuits in computers," *Elektronische Rundschau*, vol. 11, pp. 229–234; August, 1957. (In German.)
H. Gillert, "Magnetic switching circuits for the representation of logical functions," *Nachr-Tech.*, vol. 10, pp. 391–402; August, 1957. (In German.)
J. R. Bacon and G. H. Barnes, "Quantized flux counter," 1957 WESCON CONVENTION RECORD, pt. 4, pp. 246–250.
J. A. Rajchman, "A survey of the use of magnetic elements as manipulators of digital and analog information," IRE TRANS. ON CIRCUIT THEORY, vol. CT-4, pp. 210–225; September, 1957.
A. J. Meyerhoff and R. M. Tillman, "A high-speed two winding transistor-magnetic-core oscillator," IRE TRANS. ON CIRCUIT THEORY, vol. CT-4, pp. 228–236; September, 1957.
Ya. G. Koblents and D. A. Yakovenko, "Single cycle switching circuits with intermediate L-C network employing ferrites with rectangular hysteresis loops," *Elektrosvyaz'*, pp. 101–112; November, 1957. (In Russian.)
W. Matthews, "Telemetering in earth satellites," *Elec. Engrg.*, pp. 976–981; November, 1957.
D. R. Erb, "How to design a high-speed magnetic switch," *Electronics*, vol. 30.11, pp. 40–45; November, 1957.
R. W. Rochelle, "Earth satellite telemetry coding system," *Elec. Engrg.*, pp. 1062–1065; December, 1957.
R. K. Richards, "Digital Computer Components and Circuits," D. Van Nostrand Co., Inc., Princeton, N. J.; 1957.
D. H. Schaefer, "Magnetic core event counter for earth's satellite memory," *Elec. Engrg.*, vol. 77, pp. 52–56; January, 1958.
H. F. Storm, "Applications of non-linear magnetics," *Electronic Design*, pp. 32–35; February 19, 1958.
W. Matthews, R. W. Rochelle, C. B. House, R. L. Van Allen, D. H. Schaefer, and J. C. Schaffert, "Cyclops cores simplify earth-satellite circuits," *Electronics*, vol. 31.9, pp. 56–63; February 28, 1958.
H. Kihn and W. E. Barnette, "The Megacoder," RCA, Camden, N. J.; February, 1958.
K. Kandiah, "Recent developments in counting equipment," *Nuclear Instr.*, vol. 2, pp. 109–113; February, 1958.
G. Smoliar and J. Keilson, "Synchronization of a (magnetic) computer," *Electronic Design*, pp. 64–67; March 5, 1958.
D. C. Kalbfell, "Low level magnetic commutators," *Electronic Design*, pp. 70–73; March 5, 1958.
J. P. Eckert, Jr., "Transistor-Magnetic Amplifier Bistable Devices," U. S. Patent No. 2,798,169 (assigned to the Sperry-Rand Corp.), abstracted in *Electronic Design*; March 19, 1958.
A. D. Holt, "The application of square hysteresis loop materials in digital computer circuits," *Electronic Engrg.*; April, 1958.
A. D. Holt, "The application of square hysteresis loop materials in digital computer circuits," *Electronic Engrg.*, vol. 30, pp. 196–199; April, 1958.
A. Rose, "Magnetic core dividers for I.T.V. sync generators," *Electronics*, vol. 31, pp. 76–77; April 11, 1958.
"New magnetic amplifiers for digital computer uses," *Automatic Control*, pp. 52–53; May, 1958.
M. Dumaire, H. Jeudon, M. Lilamand, and J. Brodin, "Logical circuits using coils and capacitors," *Proc. Tech. Conf. on Nonlinear Magnetics and Magnetic Amplifiers*, Los Angeles, Calif.; August 6–8, 1958.
U. F. Gianola and T. H. Crowley, "Magnetic networks for performing logic," *Proc. Tech. Conf. on Nonlinear Magnetics and Magnetic Amplifiers*, Los Angeles, Calif.; August 6–8, 1958.
W. J. Dunnet and A. G. Lemack, "A new transistor-magnetic core bi-logical computer element," *Automatic Control*, pp. 32–39; September, 1958.
M. E. North and J. B. Crank, "A high-level magnetic core commutator," *Proc. 1958 Natl. Symp. on Telemetering*, paper 10.1; September, 1958.

A. Ya. Artukhin and V. Z. Kahnin, "One-cycle shift register," *Automatika i Telemekhanika*, 11 pp.; October, 1958.
H. R. Irons, "A transistor-magnetic core binary counter," PROC. IRE, vol. 46, pp. 1967–1968; December, 1958.
"Magnetic device for logic circuitry," *J. Appl. Phys.*, vol. 29, cover and page 1776; December, 1958.
M. C. Chen, "A magnetic core parallel adder," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 262–264; December, 1958.
N. S. Prywes, "Diodeless magnetic shift registers utilizing transfluxors," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 316–324; December, 1958.
R. B. Hurley, "Junction Transistor Electronics," John Wiley and Sons, Inc., New York, N. Y.; 1958 (see Chapter 22).
R. W. Rochelle, "Earth Satellite Telemetry Coding System Using Transistors and Magnetic Cores," Natl. Telemetering Conf. Rept., N.R.L., Washington, D. C.; 1958.
J. W. Kallander, "Design of minimum weight magnetic cores," N.R.L., Washington, D. C., Rept. 4895; 1958.
"Magnetic Decision Elements," Tech. Bulls., Minnesota Electronics Corp.
U. F. Gianola and T. H. Crowley, "The Laddic, a magnetic device for performing logic," *Bell Sys. Tech. J.*, vol. 38, pp. 45–72; January, 1959.
H. D. Crane, "A high-speed logic system using magnetic elements and connecting wire only," PROC. IRE, vol. 47, pp. 63–73; January, 1959. (See also *Proc. Tech. Conf. on Non-Linear Magnetics and Magnetic Amplifiers*, Los Angeles, Calif., pp. 465–482; August, 1958.)
D. C. Engelbart, "A new all-magnetic logic system using simple cores," *Digest of Tech. Papers 1959 Solid-State Circuits Conf.*, Philadelphia, Pa., pp. 66–67; February, 1959.
H. D. Crane and D. R. Bennion, "Design and analysis of MAD transfer circuitry," *Proc. WJCC*; March, 1959.
A. H. Bobeck and R. F. Fischer, "A reversible, diodeless, twistor shift register," *J. Appl. Phys.*; April 1, 1959.

16. Conference Proceedings and Books

R. N. Bozarth, "Ferromagnetism," D. Van Nostrand Co., Inc., Princeton, N. J.; 1951.
Proc. AIEE Conf. on Magnetism and Magnetic Devices—1955 (AIEE Publication No. T-78).
H. F. Storm, "Magnetic Amplifiers," John Wiley and Sons, Inc., New York, N. Y.; 1955.
Proc. AIEE Conf. on Magnetism and Magnetic Materials, Boston, Mass.; October, 1956 (AIEE Publication T-91).
Proc. AIEE Solid-State Dielectric and Magnetic Devices Conf., Washington, D. C.; April, 1957.
R. K. Richards, "Digital Computer Components and Circuits," D. Van Nostrand Co., Inc., Princeton, N. J.; 1957.
W. A. Geyger, "Magnetic Amplifier Circuits," McGraw-Hill Book Co., Inc., New York, N. Y., 2nd ed.; 1957.
Proc. Spec. Tech. Conf. on Magnetic Amplifiers (1957). Available from AIEE, New York 18, N. Y. as Publication T-98.
Proc. Tech. Conf. on Non-Linear Magnetics and Magnetic Amplifiers, Los Angeles, Calif.; August, 1958.
Proc. 4th Symp. on Magnetism and Magnetic Materials, Philadelphia, Pa., November 17–20, 1958, *J. Appl. Phys.*; suppl. to vol. 30.4; April, 1959.
M. A. Katz, "Solid State Magnetic and Dielectric Devices," John Wiley and Sons, Inc., New York, N. Y.; March, 1959.

17. Standards

AIEE Committee, "Proposed size standards for toroidal magnetic tape-wound cores," *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 76), pp. 135–137; May, 1957.
P. L. Schmidt, "Proposed size standards for toroidal magnetic tape-wound cores," *Elec. Mfg.*, vol. 60, pp. 122–124; September, 1957.
"IRE standards on static magnetic storage: definitions of terms, 1959," PROC. IRE, vol. 47, pp. 427–430; March, 1959.

18. Other Bibliographies

H. B. Rex, "Bibliography on transducers and magnetic amplifiers," *Instr.*, vol. 21, no. 4; 1948.
J. G. Miles, "Bibliography of magnetic amplifier devices and saturable reactor art," *Applications and Industry* (*Trans. AIEE*, pt. II, vol. 70), pp. 2014–2123; 1951.
"Bibliography on magnetic amplifiers for 1951–1954," *Automekhanika i Telemekhanika*, vol. 17, pp. 471–487; May, 1956. (In Russian.)
G. V. Subbotina, "Bibliography on magnetic amplifiers for 1955," *Automekhanika i Telemekhanika*, vol. 71, pp. 858–864; September, 1956. (In Russian.)
S. L. Blum, H. Zlotnick, and J. E. Zneimer, "Bibliography on ferromagnetics and ferroelectrics," *Ceram. Ind.*, vol. 67, pp. 90–92, 100–103, 116–118; 128, 130, 114–115, 1956; and vol. 68, pp. 172, 174, 176, 178, 180, 182, 184–187, 190; 1957.
G. V. Subbotina, "List of references on magnetic amplifiers and contactless magnetic elements (1956–1957)," *Automekhanika i Telemekhanika*, vol. 19, pp. 379–388; April, 1958. (In Russian.)
AIEE Committee, "Magnetic amplifier bibliography, 1951–1956", *Commun. & Electronics* (*Trans. AIEE*, pt. I, vol. 77), pp. 613–627; November, 1958.
AIEE Committee, "1957 magnetic amplifier bibliography," *Commun. & Electronics*, no. 40 (*Trans. AIEE*, pt. I, vol. 77), pp. 1051–1057; (1958) January, 1959. (Note: This is a list of 314 books, articles, and patents dealing with magnetic amplifiers. A large proportion of this work is devoted to analog control system uses, but it covers the digital field as well.)

19. Author Listing

This list is to be used to locate papers when only the author or authors are known. Beside the author's name there is a code. The first one or two characters designate the section number, the last two numbers indicate the year of publication. For example: Author S. A. Abbas, 658. To find his paper refer to the *1958* portion of section *6*.

Please note all entries within a section are in chronological order.

The Recording and Reproduction of Signals on Magnetic Medium Using Saturation-Type Recording*

J. J. MIYATA† AND R. R. HARTEL‡

Summary—This paper discusses the various factors affecting the resolution in saturation magnetic recording. The effect on the recording process of the *B-H* characteristics of the coating, coating thickness, record-head gap width, head-to-coating separation, self-demagnetization, and record-head residual magnetization are discussed. Equations are derived for the playback process relating the signal amplitude and pulse width to the coating thickness, head-coating separation, and effective gap width of the playback head. It is shown that the greatest improvement in resolution can be obtained by the development of an extremely thin coating with high ratio of coercitivity to remanence and having a rectangular *B-H* loop. The extremely thin coating will reduce the shortcomings of the record-head field pattern, the self-demagnetization effect, and the loss of resolution in the playback process.

Introduction

THE problems and limitations associated with saturation-type recording are quite different from those encountered in nonsaturation recording. In saturation recording, for the ideal case, the recording medium is completely saturated in one direction or the other. Erasure is unnecessary because the recording signal completely saturates the magnetic medium to a depth where the magnetized particles no longer influence the surface induction to a readable degree, thereby erasing all previous history. This depth on most coatings is the thickness of the magnetic medium itself.

Consider the NRZ (nonreturn to zero) recording method where the medium is saturated in one direction for a "1" and in the opposite direction for a "0". The idealized signal current or flux pattern is shown in Fig. 1(a). The playback signal using a conventional ring-type head is approximately proportional to the rate of change of the flux on the surface of the tape. Thus, the output of an ideal head when reading a flux pattern as shown in Fig. 1(a) will be a series of narrow pulses [Fig. 1(b)] corresponding to the flux changes. In general, it is from the presence or absence of these pulses that the original information is extracted. In a practical system, the actual playback signal will be composed of a series of wider pulses, as shown in Fig. 1(c). This pulse widening will seriously limit the packing density because the flux changes must be separated sufficiently so that the pulses do not interact. The interaction of the pulses may sufficiently affect their amplitude as well as their position as to give erroneous information. Fig. 2

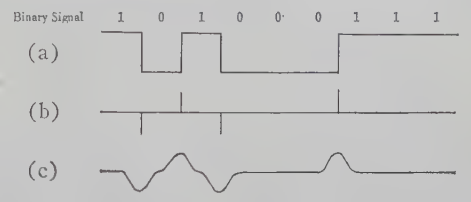


Fig. 1—Signals in saturation-type recording; a) ideal current of flux waveform (NRZ), (b) ideal playback signal, (c) actual playback signal.

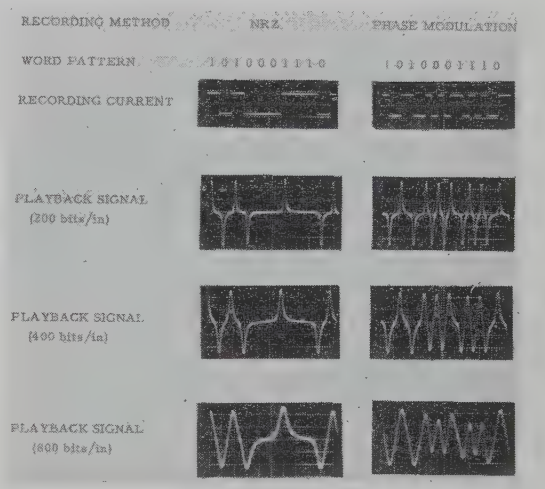


Fig. 2—Playback signals showing effect of packing density on two different recording methods (3M149 Tape).

shows the effect of pulse interaction at various packing densities using two different recording methods. It is apparent that the amplitude and pulse position may vary, depending on the information recorded. It is the sources contributing to the widening of the pulse that will be discussed in this paper. No attempt will be made to relate the pulse width to the packing density, since the latter will depend on the type of recording used (NRZ, phase modulation, etc.), the mechanical and electrical tolerances involved, and a specific definition of the conditions that constitute system failure.

There may be many sources contributing to pulse widening as seen by the block diagram of Fig. 3. The dotted boxes represent head-to-magnetic-medium separation. Since it is a relatively simple matter to eliminate many of the sources, only those that are basic to the recording and reading processes will be considered. This eliminates from consideration such factors as current and flux rise time, amplifier response, head-frequency response, and so forth. The frequency-dependent factors can be made negligible by using very low speeds.

* Manuscript received by the PGEC, September 25, 1958; revised manuscript received, March 26, 1959.
† The National Cash Register Co., Hawthorne, Calif.
‡ Clevite Electronic Corp., Cleveland, Ohio. Formerly, with The National Cash Register Co., Hawthorne, Calif.

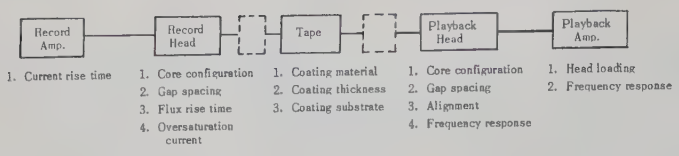


Fig. 3—Factors affecting the resolution.

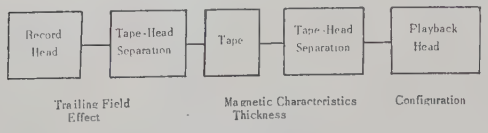


Fig. 4—Basic factors affecting resolution.

A more satisfactory breakdown is shown in Fig. 4. This paper will discuss the "wavelength"-dependent rather than frequency-dependent factors. It will be seen that the effects of the black boxes are not independent. There is an especially close interdependence between the effects of the record head and tape on the resolution obtainable.

In this paper the playback pulse width will be defined in terms of distance along the tape rather than time, because it is independent of tape speed in an ideal system. The pulse width will be defined as that distance along the tape over which the amplitude exceeds 20 per cent of the peak value. The 20 per cent level was chosen because several magnetic tape systems use this value as the clipping level. The pulse width gives an indication of the resolution capability of the system.

The actual limitation of the pulse width by the record head is dependent on the magnetic medium, the record current, and the head-to-medium separation. It is difficult to clearly divide the influences. These factors will be combined and defined as the "record-process limitation" as distinguished from the "playback-process limitation."

Recording-Process Limitation

Record-Head Trailing Effect

The record-head trailing effect is commonly defined as the demagnetization of the tape still within the field of the head as it changes polarity. To illustrate this effect, consider a typical *B-H* loop of an instrumentation tape,[1] as shown in Fig. 5(a).

From this figure, it is possible to determine the relationship between the magnetizing force H exerted by the record head and the remanence B_r of the medium. For recording purposes, it is the relationship of H to the remanence that is important. The maximum value of B_r occurs when the medium is saturated. The term retentivity (B_m) is used to denote this value. In Fig. 5(b), B_r is plotted as a function of H showing the very non-

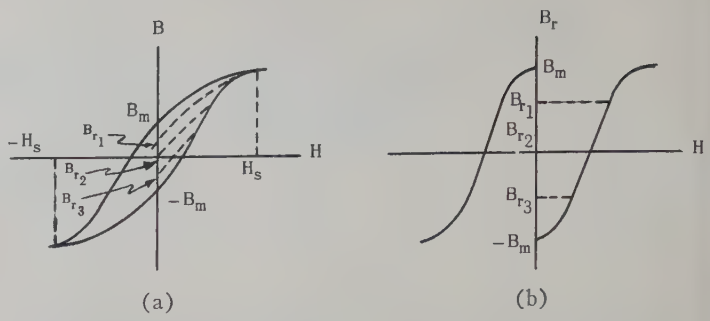


Fig. 5—*B-H* loop (a) and remanent flux curve (b) for typical oxide tape.

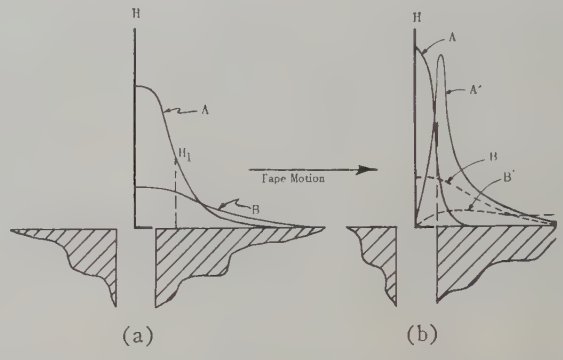


Fig. 6—(a) The magnitude and (b) the components of the field distribution (A) at the surface of a recording head and (B) some distance above the pole face. In (b) the horizontal field components are A and B, and the vertical components are A' and B'.

linear relationship. If the coating could be operated in a bistable manner (saturated in one direction or the other), this curve would not affect the final pulse shape. However, because of the flux configuration across the typical ring-type head, this condition cannot be fully achieved.

Fig. 6(a) shows the approximate distribution of the field across a typical ring recording head.[2] This field distribution is composed of horizontal and vertical components. Fig. 6(b) shows the horizontal (A, B) and vertical (A', B') field components. Both the horizontal and vertical components of the field are effective in magnetizing the tape. The relative strength of the components magnetizing the tape depends on the coating thickness and recording-field strength. The field distribution is shown only for the trailing-pole piece, since this pole exerts the final influence on the tape and cancels any effect of the leading pole. Curve A represents the field distribution along the face of the head, while curve B represents the field along a plane some distance from the head. If the record-head field changes polarity instantaneously in relation to the tape motion, all particles passing through the gap are being saturated in one direction or the other. However, a particle which has just passed the gap as the head changes polarity (position 1 with field strength H_1 of Fig. 6) will be influenced

[1] Minnesota Mining and Mfg. Co., Minneapolis, Minn., 3-M Bulletin No. 22, Tape 109A; March, 1958.

[2] S. J. Begun, "Magnetic field distribution of a ring recording head," *Audio Engrg.*, vol. 32, pp. 11–13; December, 1948.

by the new field but not saturated. Its remanence, B_r, is less than B_m. It is apparent that at the moment of flux reversal in the head, all particles under the trailing-pole face are magnetized to a varying degree dependent on the gradient across the head. The actual field pattern in the coating will depend on the permeability of the material. This permeability is not a constant but will vary from unity in the gap where the coating is saturated to a constant value at distances far from the gap. This record-head trailing effect is dependent on the *B-H* characteristics of the magnetic medium, recording current, coating thickness, record-head-to-magnetic-medium separation, and gap width of the record head.

If the field gradient across the record head had no influence on the flux pattern, one would expect the pulse width and amplitude to remain constant after saturation. However, it will be shown that they are dependent on the recording current. The increase in pulse width and decrease in amplitude may be attributed to the record-head trailing effect. The increase in pulse width due to the head-trailing effect has not been determined but it is estimated to be less than 0.2 mil for the high-resolution instrumentation tapes.

In the discussion on head-trailing effects the self-demagnetization of the magnetic medium will be neglected.

Effect of B-H characteristics: The flux distribution along the magnetic coating can be plotted from curves of Figs. 5 and 6. Fig. 7 shows the flux distribution along the coating with the flux reversal taking place when the center of the gap is at the origin, for two coatings having considerably different *B-H* characteristics. Curve (a) shows the distribution for a typical oxide coating (See Fig. 5). A coating having a more rectangular *B-H* characteristic with the same B_m and requiring the same record current (Fig. 8) will give a flux distribution as shown by curve (b). Thus one can see that considerable improvement in resolution can be obtained with magnetic coating having nearly rectangular *B-H* characteristics.

Effect of recording mmf: The effective recording point (position of maximum flux gradient) will depend on the magnetic medium as well as the recording-field strength. The relative distribution of the horizontal and vertical components of the magnetization will depend on the recording-field strength and the coating thickness. The vertical component becomes more predominant as the current is increased beyond the saturation current. Fig. 9 shows that a current greater than that required to saturate the tape (I_s) will shift the effective recording point further away from the center of the gap, reduce the amplitude (maximum rate of change of B_r), and increase the pulse width (total distance over which the flux is changing). These curves were obtained for a typical oxide coating (Fig. 5) and the record-head field distribution shown in Fig. 6, curve A. Fig. 10 shows ex-

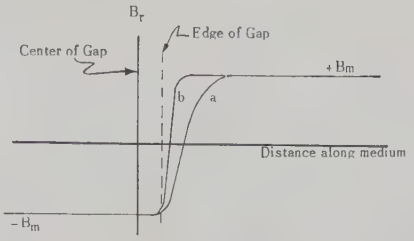


Fig. 7—Flux distribution along the magnetic medium for coatings with different *B-H* characteristics.

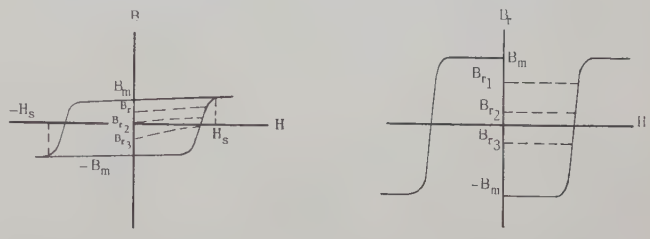


Fig. 8—Magnetic characteristics of a rectangular *B-H* loop coating.

perimental curves for the pulse width and amplitude as a function of the recording magnetomotive force. Illustrated in Fig. 11 are the shift in peak position, decrease in amplitude, and increase in pulse width for record currents 1, 2, 4, and 8 times that required to saturate the coating.

Effect of coating thickness: The record-head trailing effect becomes more pronounced as the thickness of the magnetic medium is increased, since the particles furthest away from the pole face will be subjected to a magnetic field with greatly decreased field gradient. Fig. 12 shows the flux distribution along the coating for particles in contact with the head, curve (a), and those separated from the head by the coating thickness, curve (b), as plotted from the curves of Figs. 5 and 6(a). The surface induction will be due to all distributions between curves (a) and (b), with those near the surface having the greatest influence. Curve (a) is the flux distribution with the field in the gap just large enough to saturate those particles which come in contact with the head. However, in order to saturate the particles away from the pole face, a higher record current will be required so that curve (b) or (c) of Fig. 9 would probably more accurately represent the distribution of the flux for the particles in contact with the head. The saturation current increases almost linearly with the product of coating thickness and coercivity for most of the oxide tapes (see Fig. 13).

Effect of gap width of the recording head: The gap width of the recording head will affect the signal recorded on tape insofar as the trailing-field gradient affects the flux distribution. For extremely thin magnetic coating, a very small gap would be desirable, since at the pole face surface, the field gradient will be greater

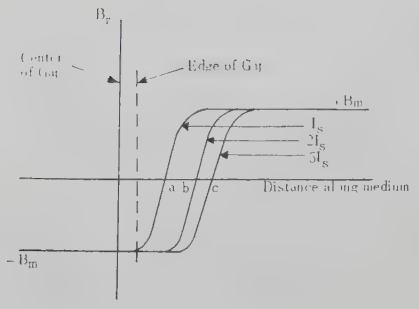


Fig. 9—Flux distribution along magnetic medium for different currents.

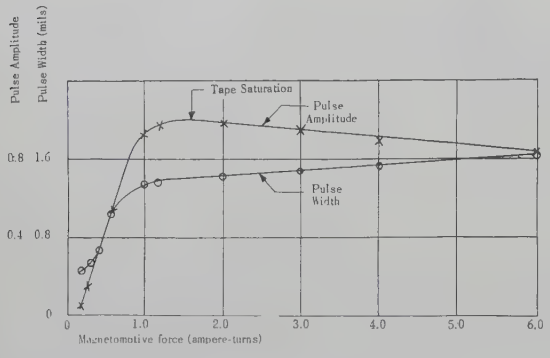


Fig. 10—Dependence of pulse width and amplitude on record current.

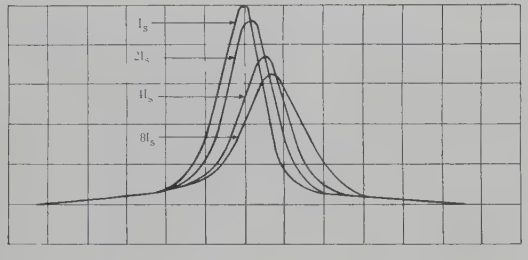


Fig. 11—Effect of record current on playback signal.

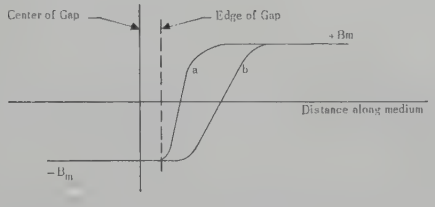


Fig. 12—Flux distribution along magnetic medium for particles in contact with the recording head (a) and those on the opposite side of the coating (b).

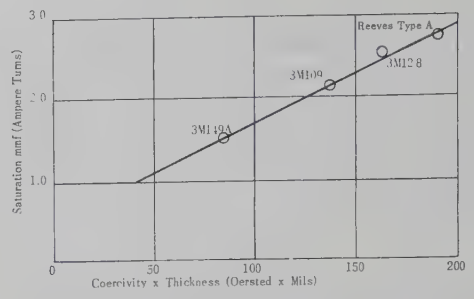


Fig. 13—Dependence of saturation mmf on product of coercivity and thickness.

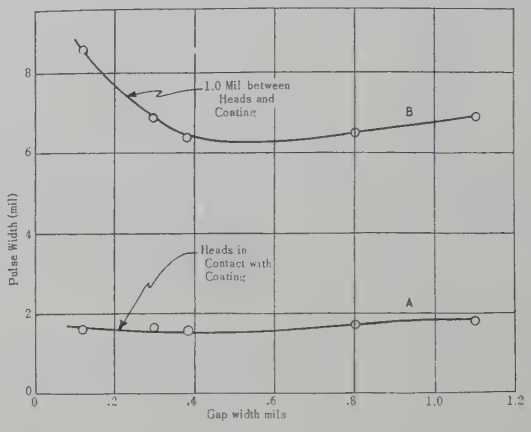


Fig. 14—Effect of record-head gap width on pulse width.

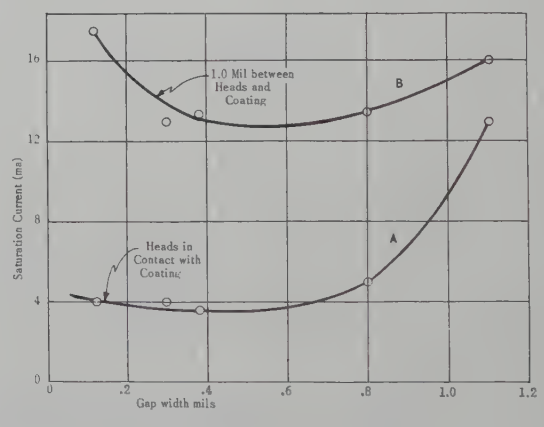


Fig. 15—Effect of record-head gap width on saturation current.

than for a wide gap. However, for relatively thick coatings, a wide gap would be desirable to set up a sufficiently strong field to saturate those particles some distance from the pole face. For very small gaps, a high mmf will be required to saturate those particles removed from the pole face, while the particles in contact with the head will be exposed to a field many times that required to saturate it. Thus, these particles will be subjected to a low effective field gradient which will result in poor resolution.

Figs. 14 and 15 show the effect of the record-head gap width on the pulse width and saturation current. The A curves were obtained with the heads in contact with the oxide coating (0.35 mil thick), while the B curves were obtained with 1.0 mil separation between the heads and the oxide. These curves show that for an in-contact system using 0.35 mil coating, the record-head gap can vary considerably before any significant change in resolution (pulse width) can occur. This is because the major limitation is in the coating and the read back process rather than in the record-head trailing field.

In systems where the coating is very thick or where there is a separation between the coating and the heads, it is important that the record-head gap be wide enough

so that the resolution will not suffer and the saturation current requirement will not be excessive.

The optimum record-head gap will depend on the coating thickness, the B-H characteristics of the magnetic material, and the separation between the record head and the coating.

Effect of record-head magnetic-medium separation: When the record head is separated from the coating, the latter is subjected to a recording field with greatly decreased gradient [Fig. 6, curve (b)]. The decreased field gradient will result in a flux distribution along the coating [Fig. 12, curve (b)] which will result in poor resolution. The loss in resolution can be reduced by using a coating with rectangular B-H characteristics. In addition, the record head should be designed to give the maximum field gradient for the head-to-coating separation being used. This will require a sophisticated design wherein the pole-face configuration is altered to give the desired distribution.[3]

Self-Demagnetization Effect

The term "self-demagnetization effect" will be applied to pulse widening caused by the field within the coating. This field will exert a demagnetizing effect upon the magnetized particles. The degree of reorientation of the particles will be dependent on the magnetic characteristics and thickness of the coating.

If the intensity of magnetization within the coating is given by I, then the field due to the magnetization will be proportional to Div I ("volume density of magnetic charge").

The field in the coating due to the volume element dv will be

$$dH = \frac{\text{Div } I\, dv}{r^2},$$

where r is the distance from the volume element dv to the point (x, y, z) in the coating.

The total field at (x, y, z) will be

$$H = \int \frac{\text{Div } I\, dv}{r^2}.$$

It is believed that this field acting on the particles within the coating will tend to reorient the particles and thereby reduce the resolution.

The loss in resolution due to self-demagnetization effect can be reduced by using a magnetic coating with a high ratio of coercivity-to-remanence to minimize reorientation of the particles. The self-demagnetization effect will be much less with material having a rectangular B-H loop. Because the demagnetizing field due to a thin coating is less than for a thick coating, the loss resolution due to this field will be smaller.

Effect of Record-Head Residual Magnetization

Since a comparatively large magnetomotive force is utilized in saturation-type recording, the residual magnetization in a poorly designed record head may be great enough to cause deterioration of the recorded signal after several passes. This effect has been especially noticeable on some of the newer thin-oxide lower-coercivity instrumentation tapes. Fig. 16 shows the effect of remanent magnetization in the record head on pulse amplitude and width after successive passes over the tape.

When the magnetized particles in the coating pass through the residual field of the record head, they are reoriented to a degree dependent upon the direction and strength of this field. The effect of this field can be shown graphically with the aid of Figs. 5 and 17. Consider a flux pattern as shown by curve (a) of Fig. 18. All points on this curve correspond to B_r of Fig. 5. Passage of a magnetized head over the coating is the same as application of a bias field $+\Delta H$ to all points of curve (a). This bias field will give a new value for the remanence B_r in the coating, curve (b), corresponding to B_r' (see Fig. 17). If the residual magnetization in the record head is now reversed, the bias field $-\Delta H$ will give the distribution shown in curve (c) corresponding to the B_r'' points.

From Fig. 18 one can see that 1) the maximum rate of change of flux and the total flux change have been reduced, 2) the rate of change of flux near the knee of the curve does not change appreciably, and 3) the total interval over which the flux is changing remains approximately the same.

Since the playback voltage is approximately proportional to the derivative of the flux curve, one can see that the residual magnetization in the record head results in a decrease in signal amplitude, and an increase in pulse width (measured at 20 per cent of peak amplitude).

The residual magnetization in the record head may lead to considerable fluctuation in playback signal. The pulse amplitude may decrease as much as 50–60 per cent in a poorly designed head. In a system using fixed level amplitude discriminators, this variation in amplitude may lead to erratic performance. The residual magnetization may be reduced by increasing the back gap or by using a core material with lower remanence.

To summarize, the greatest limitation in the recording process to high-density recording (neglecting mechanical limitations) is primarily in the magnetic medium which is intimately tied in with the record-head characteristics. By suitable choice of the coating it is possible to reduce the shortcomings in the record head. For example, Fig. 19 shows the improvement which can be obtained by using different tapes with the same record and read heads. In every case, the thinner coatings

[3] A. S. Hoagland, "Magnetic Data Recording Theory: Head Design," *Commun. and Electronics* (*Trans. AIEE*, pt. 1), paper no. 56–695, pp. 506–512; November, 1956.
———, "High-resolution magnetic recording structures," *IBM J. Res. & Dev.*, vol. 2, pp. 91–104; April, 1958.

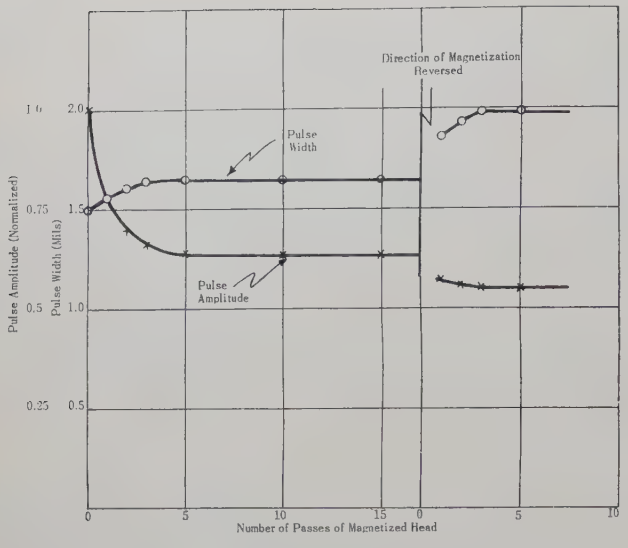


Fig. 16— Effect of residual magnetization in record head on recorded signal. Tape: 3M149A.

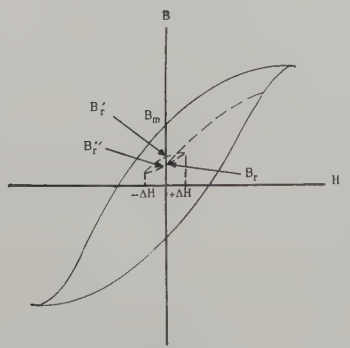


Fig. 17—Effect of record-head residual magnetization on magnetization curve.

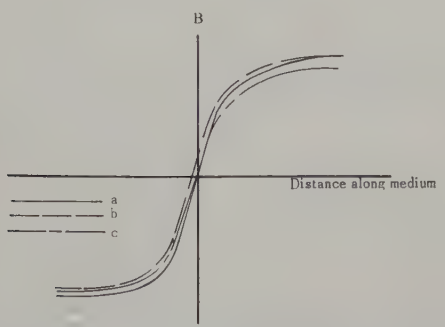


Fig. 18—Effect of record-head magnetization on recorded signal.

give better resolution. The experimental tapes have coating thicknesses of less than 0.05 mil. This improvement in resolution is believed to result from two major factors:

1) The record-head trailing effect is much less pronounced with the thinner coatings,
2) the self-demagnetizing field is much less with the thin coating.

In addition, the experimental tapes have much more rectangular *B-H* loops.

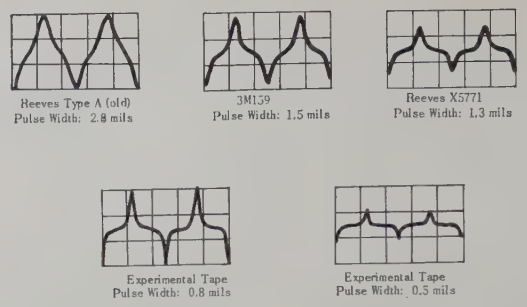


Fig. 19—Playback signals from various tapes. Recording density 300 bits/in (NRZ). Pulse width measured at lower densities.

It has not yet been determined conclusively whether the greatest limitation to high-density recording is in the record head or magnetic coating. It is true that by using thinner coatings a great improvement in resolution is obtained. Whether this is due primarily to the more favorable record-head field distribution or the smaller self-demagnetization effect has not been determined. In addition, the discussion on the playback process shows that the thinner coatings are more desirable for high resolution.

The Playback Process

One can make some calculations regarding the effects of coating thickness, head-to-medium spacing, and recording constant on the signals obtained in saturation-type recording based on theoretical calculations using idealized heads and magnetic coating. An expression will be derived showing how the signal characteristics depend on these factors. The playback signal characteristics of greatest interest will be the maximum amplitude and the pulse width measured at some clipping level. The self-demagnetization and the record-head trailing effect will be assumed to be independent of the coating thickness. The calculation of the field in the head is carried out in a manner similar to that used by Wallace.[4]

Derivation of Playback Signal

Consider the recording surface to be an infinite plate of thickness δ with the coordinate system chosen so that the central plane lies in the x-y plane as shown in Fig. 20.

To simplify the analysis, the permeability of the recording medium is assumed to be unity and the recording is purely longitudinal. Most present-day oxide tapes have a permeability of 2–3. The actual signal recorded will have both longitudinal and perpendicular components due to the recording head used. The analysis for perpendicular recording should lead to similar results.

In saturation-type recording, the ideal recording field will be as shown in Fig. 21(a). It would be desirable to have the intensity of magnetization I in the recording medium as shown by the solid curve of Fig. 21(b).

[4] R. L. Wallace, Jr., "The reproduction of magnetically-recorded signal," *Bell Sys. Tech. J.*, vol. 30, pp. 1145–1173; October, 1951.

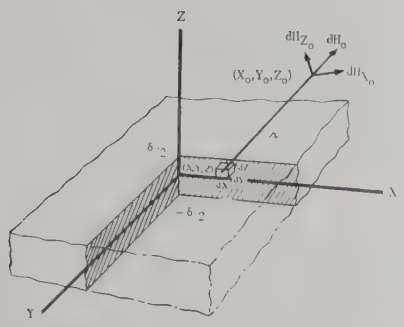


Fig. 20—Coordinate system for playback signal analysis.

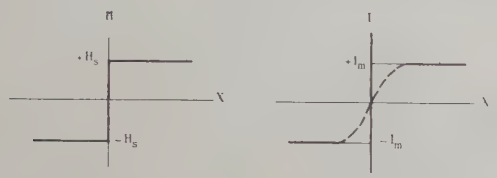


Fig. 21—Idealized record field and magnetization.

However, the actual magnetization will be as shown by the dotted curve in Fig. 21(b), because of the self-demagnetization effect in the recording medium and to the record-head trailing effect as discussed previously. The actual magnetization curve may be approximated by a function of the form arctan x/a, where the constant a determines the abruptness of the transition. The value of a will depend on the magnetic property of the medium, and, to some extent, on the thickness of the coating and the recording-head field pattern. In this analysis, the constant a will be assumed to be independent of the coating thickness and recording-head field pattern. The intensity of magnetization is assumed to be uniform throughout the depth of the medium and to have the form

$$I_x = I_m \arctan \frac{x}{a}$$

$$I_y = I_z = 0 \text{ (longitudinal recording)}.$$

By analogy with the electrostatic case, the volume density of "magnetic charge" will be

$$\rho = -\operatorname{Div} I = -\frac{dI_x}{dx} = -\frac{aI_m}{a^2 + x^2}.$$

The field at some point (x_0, y_0, z_0) due to the infinitely long element with cross section $dxdz$ at (x, z) will be

$$dH_{x_0} = \frac{\rho \cos\theta dv}{r^2} = \rho(x - x_0)dxdz \int_{-\infty}^{\infty} \frac{dy}{r^3}$$

$$= \frac{-2aI_m(x - x_0)dxdz}{(a^2 + x^2)[(x - x_0)^2 + (z - z_0)^2]}$$

$$dH_{y_0} = 0$$

$$dH_{z_0} = \frac{-2aI_m(z - z_0)dxdz}{(a^2 + x^2)[(x - x_0)^2 + (z - z_0)^2]}.$$

These are the components of the field at (x_0, y_0, z_0) due to the infinitely long element with cross section $dxdz$ at (x, z). The total field is obtained by integrating with respect to x from $-\infty$ to $+\infty$ and with respect to z from $-\delta/2$ to $+\delta/2$.

$$H_{x_0} = -2aI_m \int_{-\delta/2}^{\delta/2} \int_{-\infty}^{\infty} \frac{(x - x_0)dxdz}{(a^2 + x^2)[(x - x_0)^2 + (z - z_0)^2]}.$$

Integrate first with respect to x, using partial fractions to simplify the integration

$$H_{x_0} = -2aI_m \int_{-\delta/2}^{\delta/2} \left[\int_{-\infty}^{\infty} \frac{Ax + B}{a^2 + x^2} dx + \int_{-\infty}^{\infty} \frac{Cx + D}{(x - x_0)^2 + (z - z_0)^2} dx \right] dz$$

where

$$A = -C = \frac{(z - z_0)^2 - (x_0^2 + a^2)}{\Delta}$$

$$B = \frac{-x_0[(z - z_0)^2 + (x_0^2 + a^2)]}{\Delta}$$

$$D = \frac{-x_0[-3(z - z_0)^2 + (x_0^2 + a^2)]}{\Delta}$$

$$\Delta = (z - z_0)^4 + 2(x_0^2 - a^2)(z - z_0)^2 + (x_0^2 + a^2)^2.$$

Performing the integration and substituting gives

$$H_{x_0} = 2\pi I_m x_0 \int_{-\delta/2}^{\delta/2} \frac{dz}{(z - z_0 + a)^2 + x_0^2}.$$

Integrating with respect to z gives

$$H_{x_0} = 2\pi I_m \left[\arctan \left(\frac{z_0 + \frac{\delta}{2} - a}{x_0} \right) - \arctan \left(\frac{z_0 - \frac{\delta}{2} - a}{x_0} \right) \right]$$

$H_{x_0} = 0$, when $x_0 = 0$.

Only the x component of the field will be considered, since the z component will not contribute to the signal. For the case under consideration here,

$$z_0 \geq \frac{\delta}{2} \text{ (the field outside the coating)}.$$

For the idealized playback head, one can consider it to be a semi-infinite block of high-permeability material with the flat face spaced a distance d above the surface of the recording medium, as shown in Fig. 22. The playback signal will be proportional to the rate of change

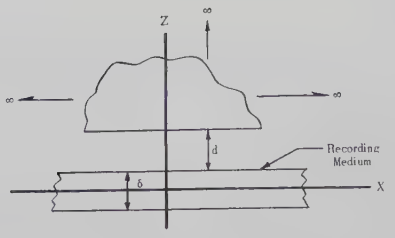


Fig. 22—Playback head position above the recording medium.

of the X component of the flux, through a plane parallel to the y-z plane as it moves uniformly in the x direction.

By using the method of images, the value of flux density in the high-permeability head can be shown to be the same as though the head filled all of space and the intensity of magnetization in the recording medium were $2\mu/(\mu+1)$ times the value actually present. Thus,

$$B_x = \frac{4\pi\mu I_m}{\mu+1}\left[\arctan\left(\frac{z+\frac{\delta}{2}-a}{x}\right) - \arctan\left(\frac{z-\frac{\delta}{2}-a}{x}\right)\right].$$

The total flux per unit width will be

$$\phi_x = \int_{d+\delta/2}^{\infty} B_x dz$$

$$= \frac{4\pi\mu I_m}{\mu+1}\left[\frac{\delta\pi}{2} - (d+\delta-a)\arctan\left(\frac{d+\delta-a}{x}\right) + (d-a)\arctan\left(\frac{d-a}{x}\right) + \frac{x}{2}\ln\frac{(d+\delta-a)^2+x^2}{(d-a)^2+x^2}\right] \quad (1)$$

for $x>0$.

And,

$$\phi_x = \frac{4\pi\mu I_m}{\mu+1}\left[-\frac{\delta\pi}{2} - (d+\delta-a)\arctan\left(\frac{d+\delta-a}{x}\right) + (d-a)\arctan\left(\frac{d-a}{x}\right) + \frac{x}{2}\ln\frac{(d+\delta-a)^2+x^2}{(d-a)^2+x^2}\right]$$

for $x<0$.

The output voltage will be proportional to the rate of change of flux ϕ_x through the plane as it moves along the magnetic medium.

$$e(x) = C\frac{d\phi_x}{dt} = \frac{2\pi\mu v C I_m}{\mu+1}\left[\ln\frac{(d+\delta-a)^2+x^2}{(d-a)^2+x^2}\right]$$

where

C is the proportionality constant
v is the velocity of the plane
a <0.

This expression relates the playback voltage e from an ideal head to the thickness of the recording medium δ, the distance between the playback head and recording medium d, and the recording constant a.

Effect of Coating Thickness and Playback Head-Coating Separation

Since only the waveshape and relative amplitude are of interest, let

$$\frac{2\pi\mu v C I_m}{\mu+1} = 1.$$

$$\therefore\ e(x) = \ln\frac{(d+\delta-a)^2+x^2}{(d-a)^2+x^2}.$$

The maximum amplitude will be

$$e_{\max} = e(0) = \ln\left(\frac{d+\delta-a}{d-a}\right)^2. \quad (2)$$

The pulse width at N per cent of peak amplitude is determined by setting

$$0.01Ne_{\max} = 0.01N\ln\left(\frac{d+\delta-a}{d-a}\right)^2 = \ln\left(\frac{d+\delta-a}{d-a}\right)^{0.02N}$$

$$= \ln\frac{(d+\delta-a)^2+x^2N}{(d-a)^2+x_N^2}.$$

x_N is the position where the amplitude has dropped to N per cent of the peak value.

$$\left(\frac{d+\delta-a}{d-a}\right)^{0.02N} = \frac{(d+\delta-a)^2+x_N^2}{(d-a)^2+x_N^2}.$$

Solving for x_N,

$$x_N = \pm\sqrt{\frac{(d+\delta-a)^2 - C_N(d-a)^2}{C_N-1}}$$

where

$$C_N = \left(\frac{d+\delta-a}{d-a}\right)^{0.02N}$$

and the pulse width is

$$= 2x_N = 2\sqrt{\frac{(d+\delta-a)^2 - C_N(d-a)^2}{C_N-1}}. \quad (3)$$

Eqs. (2) and (3) show the dependence of maximum amplitude and pulse width on coating thickness, head-to-tape separation and recording constant. The quantities d and δ can be measured directly. However, the recording constant a must be determined indirectly. Unfortunately, the constant will depend somewhat on the recording process and the coating thickness. The approximate value of a useful for the calculations can be

determined by using (3) with the signal obtained from a well-designed record head and a high-resolution playback head which are in contact with the coating ($d=0$). Table I shows the constant a as determined for various tapes using the same record-read head combination.

TABLE I

Tape	δ	Recording Constant a
*Reeves Type A (old)	0.75 mil	−0.37
3M109	0.55	−0.25
3M149	0.35	−0.2
Reeves X5771	0.30	−0.2

* This tape was received several years ago. The present Type A tape has somewhat better characteristics than this tape.

Fig. 23 shows the dependence of pulse width (20 per cent of peak) and amplitude on the playback head-tape separation using (2) and (3) with the recording constants shown in Table I. The points represent experimentally-obtained values. One can see that good agreement between theory and experiment is obtained. A small error (≈ 0.05 mil) in the measurement of the head-tape separation can account for some of the deviation from the theoretical curves. These curves show what would happen to the playback signal in a drum system where relatively large head-coating separation is employed. However, it should be noted that in such a system, the recording constant a would vary with the separation because of the different effective field gradient of the record head.

The effect of coating thickness on the playback signal can be determined from (2) and (3). Fig. 24 shows the calculated pulse width and amplitude for recording constant equal to −0.2 (3M149), assuming this constant to be independent of the coating thickness. It was determined for a coating thickness of 0.35 mil. It includes the effect of self-demagnetization and record-head trailing field for this thickness. Since these effects are a function of the coating thickness, one should expect the actual curves to deviate from these curves for thicknesses greater and less than 0.35 mil. Since coatings were not available, with the same magnetic properties but different thicknesses, the theoretical curves could not be verified. Practically, the pulse-width curve should lie above the theoretical curve for thicknesses greater than 0.35 mil and below it for thinner coatings because of the dependence of a constant a on the coating thickness. The converse is true for the amplitude curve. The difference, if any, between the theoretical and actual curves may be attributed to the self-demagnetization and record-head trailing effects which are dependent on the coating thickness.

Effect of Gap Width of the Playback Head

The analysis in the previous section was based on an idealized playback head with infinitely narrow gap so

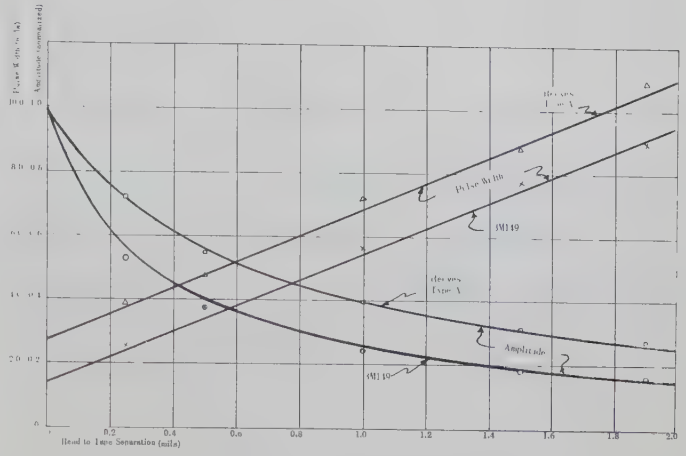


Fig. 23—Pulse width and amplitude as a function of separation between readback head and tape.

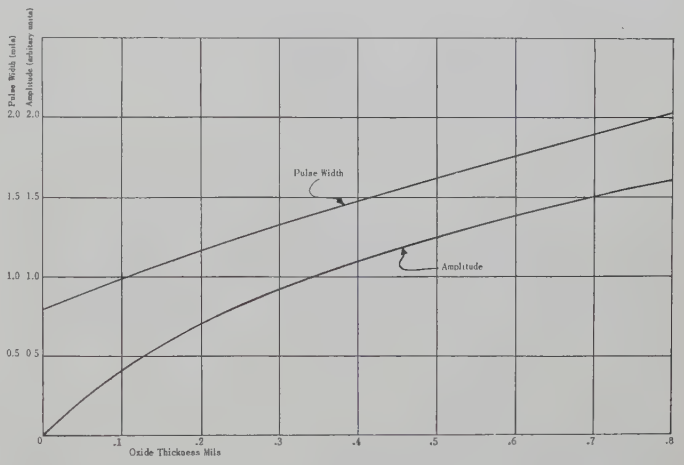


Fig. 24—Pulse width and amplitude as a function of tape thickness. Tape constant $a=-0.2$.

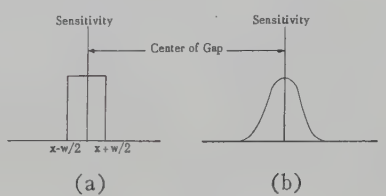


Fig. 25—(a) Idealized playback head sensitivity contour. (b) Realistic playback head sensitivity contour.

that the playback voltage was assumed to be proportional to the rate of change of flux through a plane parallel to the y-z plane. Consider now a playback head with a sensitivity contour as shown in Fig. 25(a).

All flux lines in the interval $x-w/2$ to $x+w/2$ pass through the coil in the head, where they are effective in producing a voltage. The playback voltage will be proportional to the rate of change of flux through the coil. The total flux intercepted by the head will be

$$\phi_x = \int_{x-w/2}^{x+w/2} \phi_x dx.$$

The rate of change of flux will be

$$e_x = k \frac{d}{dx} \int_{x-w/2}^{x+w/2} \phi_x dx \tag{4}$$

where k is the proportionality constant depending on the number of turns in the head, tape speed, and so forth. In any practical head it will also be dependent on w, since it is a function of the reluctance across the gap and through the core. However, since it is a multiplying factor, the general waveshape will not be affected. The factor k, therefore, will not affect the pulse width.

Using the formula of Leibnitz, (4) becomes

$$e_x = k[\phi_{x+w/2} - \phi_{x-w/2}].$$

Thus, the playback voltage with the head at position x is proportional to the difference between the flux through the plane at $x-w/2$ and $x+w/2$.

The maximum signal amplitude is

$$e_{\max} = e_0 = k[\phi_{w/2} - \phi_{-w/2}]$$

but

$$\phi_{w/2} = -\phi_{-w/2}.$$

$$\therefore\ e_{\max} = 2k\phi_{w/2}.$$

For pulse width at N per cent of the peak amplitude, the following equation must hold

$$0.02N\phi_{w/2} = \phi_{x_N+W/2} - \phi_{x_N-w/2}. \tag{5}$$

x_N is the head position where the amplitude has dropped to N per cent of the peak value.

The solution of this equation for $2X_N$ will give pulse width as a function of the effective gap width. Eq. (5) cannot easily be solved directly for X_N, because ϕ is a transcendental function. The equation therefore was solved graphically. The method used to calculate the pulse width $2X_N$ will be illustrated for an in-contact system using a playback head with 0.2-mil effective gap width and 3M149-A instrumentation tape. For this system, the following parameters are used

$$\delta = 0.35 \text{ mil}$$

$$d = 0$$

$$a = -0.2$$

$$N = 20\%$$

$$w = 0.2 \text{ mil}.$$

Substituting these values in (1) and (5) gives

$$0.4\phi_{0.1} = \phi_{x_{20}+w/2} - \phi_{x_{20}-w/2}.$$

The left side of this equation is equal to 0.039. The value of the right can be solved for various values of x. The value of x_{20} satisfying the equation will be the intersection of the $\phi_{x_{20}+w/2} - \phi_{x_{20}-w/2}$ curve with the straight line $0.4\phi_{0.1}$. (See Fig. 26.)

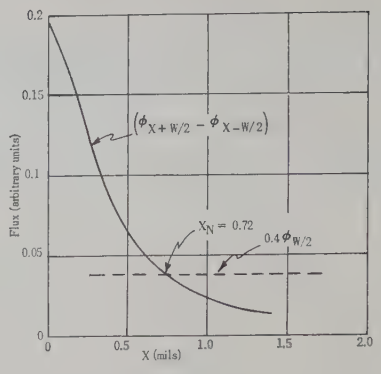


Fig. 26—Graphical solution of (5) for X_N using a playback head with 0.2 mil effective gap width.

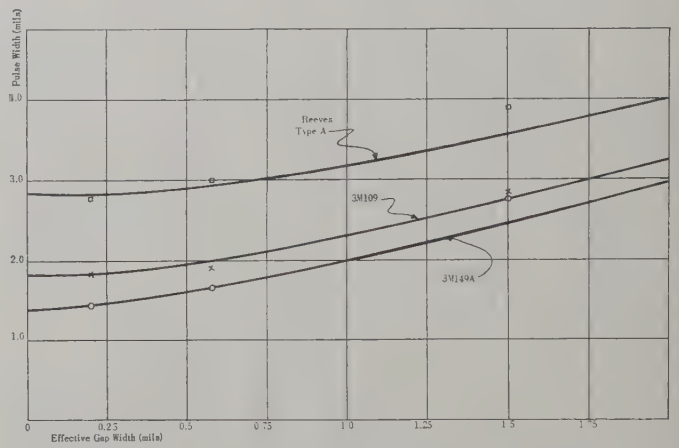


Fig. 27—Pulse width as a function of effective gap width for various tapes.

Fig. 27 shows the calculated dependence of pulse width on the effective gap width of the playback head. These curves were calculated for the case where the head was in contact with the medium ($d=0$), using (1) and (5).

In any conventional ring-type head, the sensitivity curve will not be as sharply defined as shown in Fig. 25(a). Fig. 25(b) more nearly represents the actual sensitivity. The effective gap width is not necessarily equal to the physical gap width. The experimental points for Fig. 27 were obtained using the effective gap width, as determined by measuring the wavelength at which the first minima in signal was observed.[5] The points were obtained using heads having similar configuration but different gap spacings.

From the curves of Fig. 27, one can see what effective gap widths must be used. For example, if the pulse width is not to exceed 10 per cent of the absolute minimum, the effective gap widths of the playback head

[5] S. J. Begun, "Magnetic Recording," Rinehart & Co., New York, N. Y., p. 85; 1955.

must not exceed 0.4, 0.5, and 0.7 mil for the 3M149, 3M109, and Reeves Type A (old) tapes, respectively. In general, it would be desirable to use the largest gap width consistent with satisfactory performance. The larger gap width will simplify the manufacture of the heads. For extremely small gaps, small differences in gap length from head-to-head will greatly change the ratio of the reluctance of the front gap to that of the path through the core, causing considerable variations in signal amplitudes.

Summary

From the analysis presented, one can see that for highest resolution, the following objectives should be sought:

1) Magnetic coating with rectangular *B-H* loop to reduce the record-head trailing effect and the self-demagnetization effect.
2) High ratio of coercivity to remanence to reduce the self-demagnetization effect.
3) Minimum record current, consistent with satisfactory operation, to reduce the record-head trailing effect.
4) Minimum coating thickness, consistent with satisfactory operation, to reduce the record-head trailing effect, self-demagnetization effect, and loss of resolution in the playback process.
5) Minimum separation between head and coating to reduce record-head trailing effect and loss of resolution in the playback process.

The maximum effective gap width that can be used will depend on the coating, record head, and head-to-coating separation. An indication of the behavior to be expected can be calculated using the equations derived in this report. In general, it would be desirable to use the largest gap width which will give the required resolution so that the tolerances involved in construction of the heads can be relaxed and the performance characteristics made more uniform.

Conclusions

The basic factors limiting the resolution in saturation-type recording has been discussed and equations derived for the playback process which relate the signals to some of the variables in the system. This analysis, based on certain simplifying assumptions, leads to a better understanding of the over-all performance to be expected in any system.

It has been shown that the greatest improvement in resolution can be obtained by the development of an extremely thin coating with a high ratio of coercivity to remanence and having a rectangular *B-H* loop. This coating would greatly reduce the shortcomings in the record-head field pattern.

Magnetic Core Logic in a High-Speed Card-to-Tape Converter*

E. BLOCH† and R. C. PAULSEN†

***Summary*—This report describes a static magnetic shift circuit and the logical connectives derived from it. The prime advantages of magnetic circuits are their low cost, high reliability, and ease of maintenance. The application of these circuits to the design of a card-to-tape converter is discussed.**

I. Introduction

MAGNETIC devices exhibiting square hysteresis loop characteristics have proved useful and extremely reliable computer memory elements. The low cost and practical small size of these devices have made magnetic cores the standard for use in the high-speed, random-access, large-capacity working storage of a computer. However, it is of historic interest to note that the first suggested use for such devices was in the construction of static delay lines or shift register circuits.[1] This report describes such a static magnetic shift circuit and the logical connectives derived from it. In addition, the application of these circuits to the design of a card-to-tape converter is presented. These results arose from a research program initiated to in-

* Manuscript received by the IRE, November 18, 1958; revised manuscript received March 31, 1959.
† IBM Product Dev. Lab., Poughkeepsie, N. Y.

[1] A. Wang and W. D. Woo, "Static magnetic storage and delay line," *J. Appl. Phys.*, vol. 21, pp. 146–149; January, 1950.

vestigate magnetic logic-performing circuits with the goal of developing a family of compatible, low-cost, highly reliable, logical connectives. The circuits developed by this program and a description of their operation are discussed.

Upon completion of the circuit development effort, a need for further magnetic circuit research where machine motivation would be present was felt to exist. Typical postulated problems to be solved by the construction of a model were: serviceability of pulse logic systems; power distribution problems encountered in circulating the current pulses required with magnetic circuit logic of the type described; system flexibility in relation to ease of modification of the machine design for logic changes, added features, etc.; an investigation of many system permutations of the basic building blocks to ensure maximum compatibility of interconnection of the blocks; and definite establishment of machine types where use of the proposed circuits would be advantageous.

The machine chosen for modeling was the IBM 714/759 Card Reader-Synchronizer. This machine performs card-to-tape code conversion and speed synchronization and also processes the asynchronous card information to the completely synchronous computer central processing unit. A vacuum tube-germanium diode version of this machine is available for comparison purposes, and the internal operating speed of the tube circuits is consistent with the operating speed of the magnetic logic circuits. A description of the model system as implemented with magnetic logic, including pulse driver design and the circuit packaging utilized, is given. A component-count comparison between the data-flow model and the extant tube-diode version has been tabulated. Of particular interest is the reduced number of cams and relays used in the model.

II. Device Characteristics

The device used in all the circuits of the card-to-tape converter is a ferrite toroidal core with the following dimensions:

I.D. 0.080″
O.D. 0.115″ } Diametral ratio: 1.44.
Height 0.055″

The material is a standard ferrite of the type used in many of IBM's large computer memories. Typical characteristics are:

Retentivity, B_r	2000 gauss
Saturation Flux Density, B_s	2200 gauss
Switching Coefficient, S_w	1.0 oersted-μsec
Coercive Force, H_c	1.1 oersted.

The selection of physical size for the toroidal core when used in logical building blocks is by necessity a compromise between several factors, including

a) ease of wiring and assembly,
b) minimum drive current requirement,
c) minimum drive power requirement, and
d) optimum signal levels in operation.

Factors a) and b) exert opposite influences on the selection of the diameters. The same is true of c) and d) on the selection of the height of the core.

With regard to electrical characteristics, it is important that the material have a good B_r/B_s and a good squareness ratio. Squareness ratio is defined as the ratio of ampere turns required for saturation to those required to drive the core to the first break in the pulsed flux curve (see Fig. 1). The squareness ratio is 1.65 for a diametral ratio of 1.44. A convenient figure of merit M may be stated by the following definition:

$$M = \frac{\dfrac{NI \text{ Saturation}}{NI \text{ Threshold}}}{\text{Diametral Ratio}} = \frac{\text{Squareness Ratio}}{\text{Diametral Ratio}}.$$

For the core under discussion, M is equal to 1.14 with 1.0 as an ideal value of M. The high threshold mmf allows the use of preferential dc bias in the operation of the logical connectives.

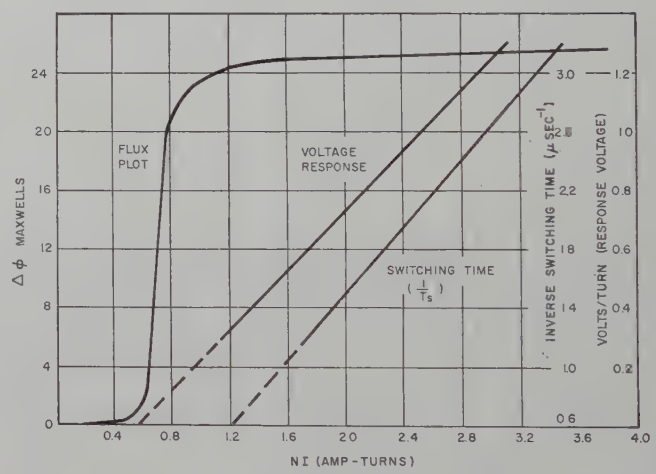


Fig. 1—S-curve, peak-voltage response, and switching curve vs net ampere-turns.

Fig. 1 also shows plots of inverse switching time and peak voltage response (all vs net ampere-turns). To select cores that are suitable for logical circuit operation, two drive currents were chosen: one at the threshold and the other well into saturation. Both flux changes and switching times are observed and must fall within close limits. The core must also pass a disturb test to demonstrate its stability and the squareness of minor loops. Since this test is usually performed when the ferrite core material is first selected for memory operation, it is not described in detail in this report.

III. Core Logic Fundamentals

All core logic circuits reported in the literature make use of certain device characteristics and circuit techniques. Some of the obvious device characteristics utilized are: the bistability of the toroidal core to provide a storage or memory function, the pronounced nonlinearity of the flux plot to permit coincidence selection, and the controlled switching constants and flux reversals. Conceptually, the toroidal core can be regarded as either a transformer or a variable impedance. Considered as a transformer, the core will deliver an output under drive excitation if previous excitations have left it in the set state. If the core has been left in the reset state, no such output occurs. On the other hand, the core, when viewed as a variable impedance, presents a high impedance to the drive source if driven to the reset state from the set state or vice versa. A negligible impedance is presented to the drive source if the core is shuttled from a remanence point to the corresponding saturation point.

Certain techniques, other than those mentioned above, are used in core logic to produce the required functions. To explain the first technique, it should be noted that the summation of drive ampere-turns applied to a particular core [see Fig. 2(a)] is expressed by

$$(NI)_{\text{total}} = \sum_{i=1}^{n} (NI)_i.$$

If the arithmetic sum of the drive ampere-turns exceeds the threshold, the core will switch; if not, the core will not switch.

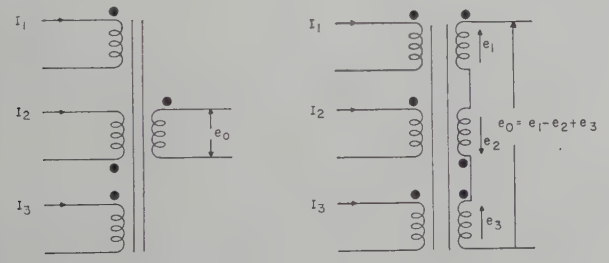


Fig. 2—Core logic techniques.

A second technique employs the summation of induced voltage on series connected secondaries [see Fig. 2(b)]. The induced voltage, e_0, can be applied to a winding on a second core. The resulting flux change $\Delta\phi$ in the second core depends on the magnitude of the output voltage e_0 and the time Δt for which it is applied. This is given by Faraday's Law

$$\Delta\phi = \int_{t_0}^{t} e_0 dt.$$

Also related to the discussion of techniques is the ability to choose an appropriate impedance scale for core circuits, depending on their application. This property is unique since it does not entail a new circuit design but strictly a change in geometry of the core and/or a change in winding turns.

The operation of the circuits is sequential: input signals are applied during one-half clock cycle and output signals are delivered one-half clock cycle later. This method of operation is inherent in this type of core logic, and proper phasing of circuits must be observed in the logical construction of an over-all system. This will be discussed in greater detail subsequently.

An important factor in the operation of the circuits is the signal gain. This gain can be defined as the ratio of the output power to the input power which is applied one-half clock cycle earlier. (Signal gain may also be expressed in terms of energy (w) and flux changes ($\Delta\phi$) with equal convenience.) The output power for each logical connective is sufficient to operate the inputs of as many as three following circuits which eliminates the need for active interstage amplification.

The over-all circuit gain is attributed to several factors. A principal factor is that losses due to the high threshold mmf of the core are minimized by the use of a dc biasing current because the mmf supplied to the core from this source is just sufficient to bias the core to the threshold.

As a core switches, the output current is proportional to the drive current. Since the drive current can be made arbitrarily large, the current (I), which flows in the output circuit, can be adjusted as required. The energy (w) delivered to a load circuit is given by

$$w \approx I\Delta\Phi,$$

where $\Delta\Phi$ is a constant. The input signal energy occurring during a previous clock phase may be relatively small as compared with the output energy delivered in the succeeding clock phase. It should be realized that the gain obtained in this manner is at the sacrifice of time. This time delay is analogous to the half-cycle delay in magnetic amplifiers.

IV. Circuit Description

A set of logical connectives must fulfill certain requirements if it is to be useful in the design of computing machines and logical networks. The logic circuits must be compatible with each other in order to permit their random interconnection. The circuits must be designed to permit branching with a high degree of isolation between the driving and driven circuits. This latter requirement is especially important in the circuits under discussion, since transformer coupling exists between stages of logic. A further and rather obvious requirement is stability of circuit operation with varying parameters such as drive currents, resistor values, and frequency of operation.

In contrast to diode and vacuum tube logic, magnetic device logic lends itself to the design of logical connec-

tives other than AND, OR, and Inversion. Therefore, a judicious choice of logical functions must be made so that the list is not prohibitively long. Before discussing the building blocks, the basic transfer circuit will be described, since all logical connectives are variations of it.

A. Transfer Circuit

This transfer circuit is based on the shift register described by Russell.[2] As is common with core circuits, the output of the transfer circuit is a pulse delayed by one-half clock cycle from the input pulse. If transfer circuits are connected in series, a shift register results which is operated in an (*A*, *B*) or double-rank fashion.

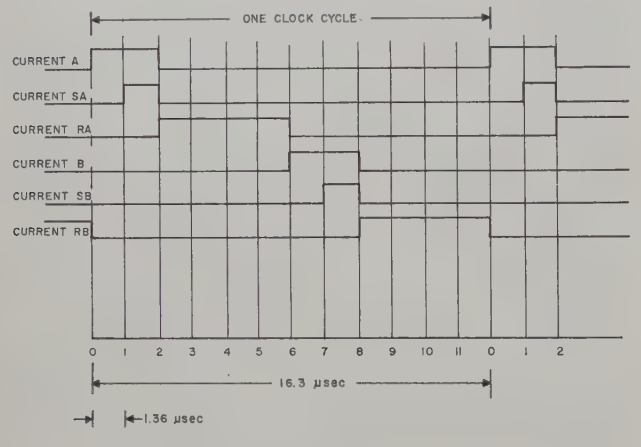


Fig. 3—Pulse schedule.

The clock or drive current sequence is shown in Fig. 3.

The period of 16.3 μsec was chosen for use in the Card-to-Tape Converter model because this period is an integral submultiple of the respective periods of the 727 Tape Unit and the 705 Central Processing Unit—the units with which the model will operate. Although the choice of a convenient period is usually left to the machine designer, it should be borne in mind that core heating, which is caused by internal switching losses, limits the described circuits to a lower limit of 5.0 μsec. Improved core materials or cooling techniques would permit reduction of the period to a new limit established by circuit constants.

Only the signal wiring will be shown, particularly the input and output circuit and the coupling or storage loop. Drive windings and the pulse schedule of drive windings are indicated by capital letters (*A*, *SA*, *RA*, etc., in Fig. 4). The 1 or 0 in parentheses following the letter designation indicates the direction of the current drive. In the "1" case the current drives the core to the one or set state; the "0" case refers to a drive to the zero or reset state. DC refers to a dc bias current of constant amplitude applied to the core. The bias mmf produced is slightly less than the threshold of the core and is normally to the ONE state. This bias is used to minimize the mmf that must be supplied from signal sources to switch the core.

Loop windings are shown with polarity markings. Current into the dotted end of a winding resets the core to the zero remanence state inducing a voltage on any winding of the same core. The voltage induced will be such as to make the dotted end of the winding positive as compared to the undotted end.

The drivers used in conjunction with the circuits described must approximate constant current generators in order to accommodate large variations in load, since a variable number of cores switch depending on the information to be processed. Although high-power peaks are encountered during the time the cores are switching, the average power dissipation is low. At the present time thought is being given to a solid-state driver element. However, suitable devices are not commercially available in quantities. The current generators used for these circuits therefore employ pentode vacuum tubes. A description of this driver will be given subsequently.

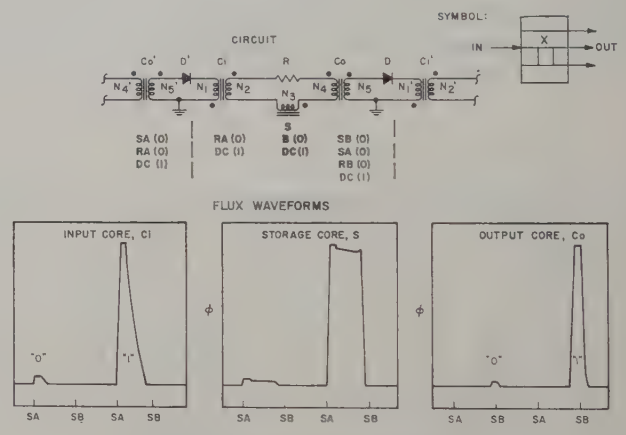


Fig. 4—Function: transfer.

The transfer circuit is shown in Fig. 4 in its normal environment; *i.e.*, the output coupling core, C_o', of the preceding stage and the input coupling core C_i' of the following stage are shown connected to the circuit. A flux diagram for each of the three cores of the transfer circuit illustrates the status of the cores for both a ZERO and a ONE input. A zero input is represented by the absence of an input pulse. All cores remain in the zero state throughout the entire cycle, and the zero output is represented by the absence of an output pulse. Since no significant changes occur for a ZERO transfer, the operation described below is for the transfer of a ONE bit.

Assume that the C_o' core is in the ONE or set state and that all other cores are in the ZERO or reset state. Input occurs when the C_o' core is reset at *SA* time. The induced voltage across the N_5' winding causes current to flow through diode D' into the undotted end of the N_1 winding on the C_i core.

[2] L. A. Russell, "Diodeless Magnetic Core Logical Circuits," IBM Res. Rept. No. RC-2; March 19, 1957.

The input current will set the C_i core, inducing a voltage across the N_2 winding with the dotted end negative. The resulting counterclockwise loop current flows into the undotted end of the storage core S, setting this core. The output core C_o is prevented from switching during input time by means of a HOLD winding SA (0). Thus, the input current drive causes the concurrent switching of both the C_i and S cores to the ONE state. After the input drive terminates, these two cores are in the set state, and all others are in the reset state.

The next drive pulse, according to the sequence of Fig. 3, is the RA drive which is used to reset the input coupling core C_i. When this core is reset, voltages will be induced in the loops, including windings N_1 and N_2.

Consider first the loop including N_1. As the C_i core is reset, a clockwise loop current flows. This current tends to set the C_o' core in the preceding circuit. However, this action is prevented by applying a hold winding to the core RA (0). The output voltage of N_1 is dropped across the forward resistance of the diode D' and the wire resistance of this loop.

Consider now the loop including winding N_2. A clockwise current will flow into the dotted ends of windings N_4 and N_3. The C_o core is in the reset state, and there is no significant change for this core. The S core is in the set state, however, and tends to be reset by this current. Because this core is dc biased to the ONE state, an appreciable current can flow in this loop before the S core will start to reset. This loop current can be adjusted to less than the critical value by resetting the C_i core at a relatively slow rate and by including a series resistance, R. Thus, the C_i core is reset without affecting the status of any other core. The S core remains in the set state while all others remain in the reset state.

The output operation is initiated by resetting the S core with the B (0) drive. As this core resets, the induced voltage across N_3 causes a counterclockwise loop current to flow. Current flow through N_2 is of no significance since this core is already in the reset state. Current flow through N_4, however, sets the C_o core. Current flow in the N_5 circuit is blocked at this time through the action of diode D.

The C_o core is employed as an output power amplifier. Although Fig. 4 shows a single input core C_i' as the load, the C_o core provides sufficient output drive for up to three succeeding circuits. The additional circuits would be connected in series with the N_1' winding.

The output pulse is delivered to the load at SB time. The SB drive resets the C_o core, inducing voltages across the N_4 and N_5 windings with the dotted ends positive. The loop current which flows in the N_4 circuit cannot erroneously set the S core because the B drive is still applied and holds this core in the reset state. The induced voltage across N_5 causes current to flow through diode D into the N_1' load circuit. This corresponds to the N_1 input current previously described except that it occurs at SB time. The entire operation for this next

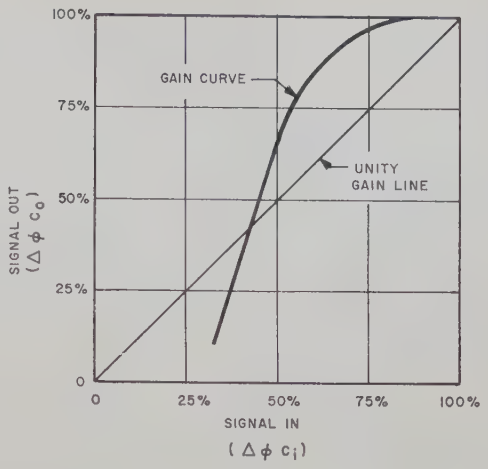


Fig. 5—Gain curve.

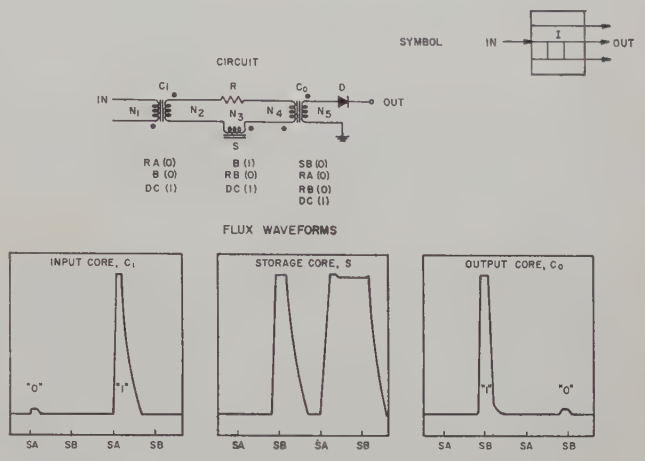


Fig. 6—Function: inverter.

circuit is just as described above but with the alternate clock phase.

A typical gain curve for the transfer circuit is shown in Fig. 5. Input core vs output core flux changes are plotted. Although no active components are used, the circuit itself provides gain as described previously.

B. Inverter Circuit

The inverter circuit is similar to the transfer circuit and is shown in Fig. 6. Inversion is obtained by modifying the output mode of operation of the transfer circuit. The significant changes made are the reversal of the B drive, B (1), and the reversal of the N_4 winding polarity. Other changes made in drives are described subsequently.

The technique used for inversion can be applied to any positive connective derived from the basic transfer circuit to obtain the corresponding negative function. For example, the OR function becomes $\overline{\text{OR}}$; the AND function becomes $\overline{\text{AND}}$, etc.

The operation will be described for both a ZERO and a ONE input. A flux diagram for each of the three cores illustrates the flux status for both input conditions.

Assume that a ZERO input occurs. At input time both the C_i and the S cores will remain in the reset state. At B time the S core is set by the B (1) drive. The induced voltage across N_3 causes a clockwise current to flow into the undotted ends of both N_2 and N_4. A B (0) winding has been added to the C_i core to prevent this core from setting at this time. The C_o core, however, is set to the ONE state. At SB time the C_o core is reset, and a ONE output pulse is delivered to the load. At RB time the S core is reset. The induced voltage across the N_3 winding during reset is dropped across the resistor R.

Next, suppose that a ONE input occurs. At input time both the C_i and the S core will be set. Next the C_i core is reset leaving the S core in the set state. Because the N_4 polarity was reversed, the C_o core is held in the reset state at RA time rather than at SA time. When the drive pulse B (1) occurs, the S core is already in the set state and no significant voltage is induced across the N_3 winding. Accordingly, the C_o core remains in the reset state. Thus, when SB occurs, the C_o core is already reset and no current or a ZERO is delivered to the load. The S core is reset at RB time as before.

C. OR Circuit

An inclusive OR circuit is shown in Fig. 7(a). Two more input coupling cores are added to the basic transfer circuit. Their secondary windings are series connected in the circuit loop. Setting any one or more of the input cores by a signal input pulse will induce a voltage, the voltage producing a loop current which sets the S core. Subsequent operations are identical with those of the transfer circuit.

The maximum number of input circuits which can be added is determined primarily by the signal-to-noise ratio of the inputs. Since the zero inputs are additive, the signal-to-noise ratio will decrease as more input circuits are added. With three inputs, however, the noise level is still below the point where circuit gain occurs, and the normal signal-to-noise ratio is preserved. A choice of three inputs was supported by a systems study which indicated that this number would be adequate for the majority of logical situations.

D. NOT OR Circuit

The NOT OR circuit is shown in Fig. 7(b). An output is obtained only when all three inputs are zero. The inverter mode of operation is applied to the OR circuit to obtain this function.

E. AND Circuit

The AND circuit is shown in Fig. 8(a). Three separate signal input windings are applied to the input coupling core C_i. An additional current drive winding SA(0) drives the C_i core to the reset state at input time. The magnitude of this drive SA(0), is adjusted so that the

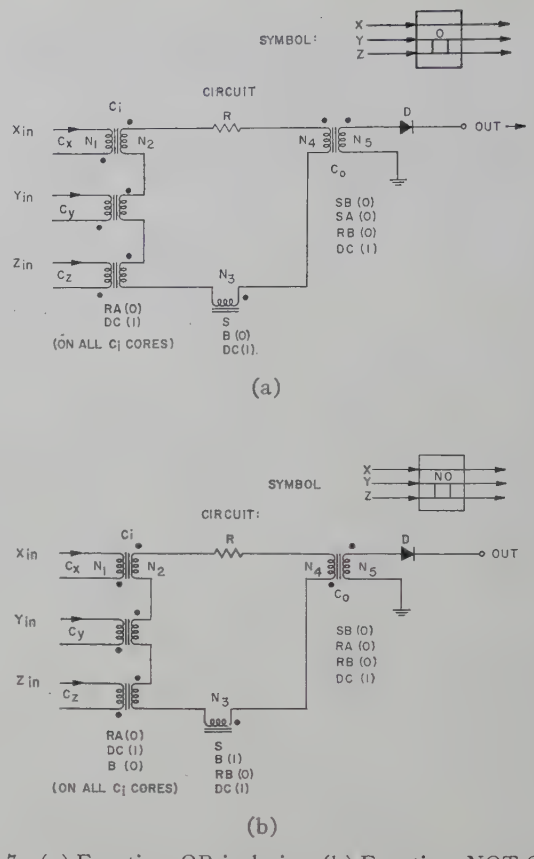


Fig. 7—(a) Function: OR inclusive. (b) Function: NOT OR.

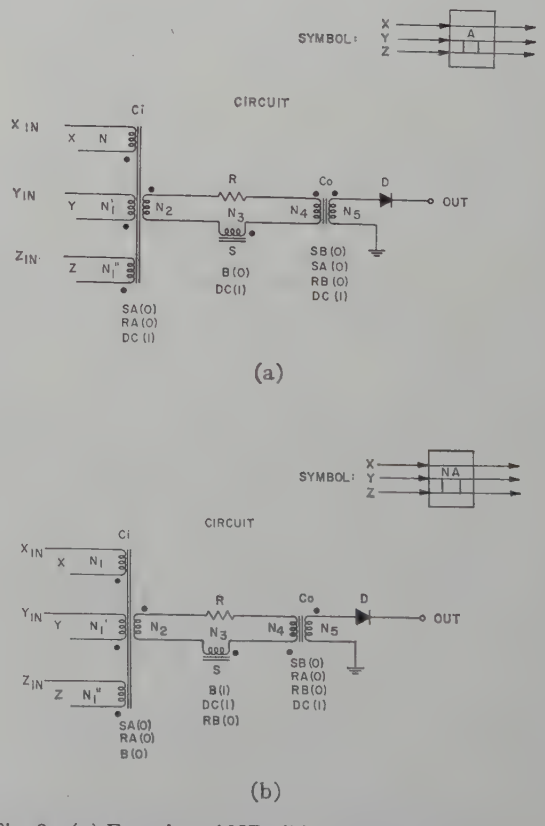


Fig. 8—(a) Function: AND. (b) Function: NOT AND.

mmf which drives the core to the reset state is equal to the sum of two input mmf's which drive the core to the set state. Thus, all three inputs must occur in order to set the C_i core. (A two-way AND circuit can be obtained by adjusting the $SA(0)$ drive to equal one unit input mmf.)

Once the C_i core is switched, *i.e.*, all inputs present, the operation is identical with that described for the transfer circuit.

The AND circuit is limited to three inputs for the same reasons stated in Section IV(C). In addition, variations in signal levels tend to make discrimination more difficult as the number of inputs is increased. For example, a six-way AND circuit will result in a discrimination ratio of 5 to 6. It is obvious that this is hardly a practical condition.

F. NOT AND Circuit

The NOT AND circuit is shown in Fig. 8(b). The circuit configuration is similar to the AND circuit. An output is obtained from the C_o core when one or more inputs are zero. The inverter output mode of operation is applied to the AND circuit previously described.

G. $X \cdot \overline{Y}$ Circuit

The $X \cdot \overline{Y}$ circuit, shown in Fig. 9(a), uses two input coupling cores. Their loop windings are connected so that the induced voltages will be of opposite polarity.

With an input to C_x only, the voltage induced across the N_2 winding sets the S core. Normal transfer from S to C_o results in an output for the given input condition.

The N_6 winding is a single turn on the C_y core which tends to inhibit the switching of the C_y core by the counterclockwise loop current when only the X input is present.

With C_x and C_y inputs both present, the N_2, N_2' voltages cancel and the storage core is not set at input time.

With both inputs zero, the storage core again is not set at input time; therefore no output is obtained.

With only a Y input present, an output is induced across N_2' which switches C_x to the one state. The loop current is into the dotted end of the winding on the storage core; hence it is not switched. At reset time the outputs from C_x and C_y oppose each other.

H. $\overline{X}+Y$ Circuit

The $\overline{X}+Y$ circuit is shown in Fig. 9(b). The circuit configuration is identical with the $X \cdot \overline{Y}$ circuit. An output occurs whenever an input to C_y occurs and/or whenever an input to C_x does not occur. The inverter output mode of operation is applied to the $X \cdot \overline{Y}$ circuit described above.

I. Circuit Parameters

The turns used for all circuits in the feasibility model were

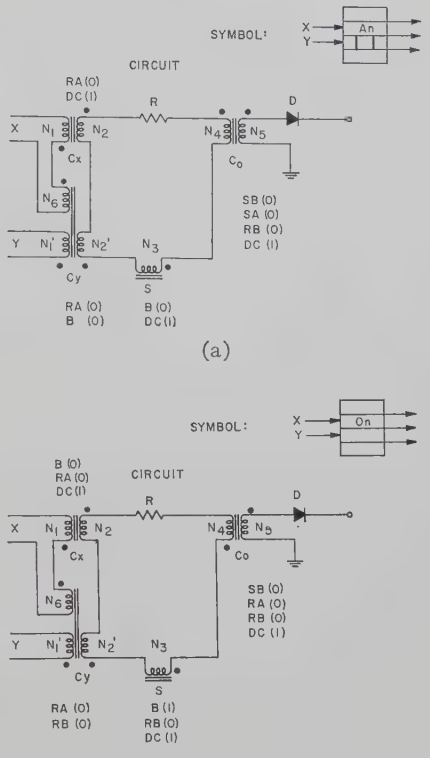


Fig. 9—(a) Function $X \cdot \bar{Y}$. (b) Function: $\bar{X}+Y$.

$N_1 = 5$ turns
$N_2 = 12$ turns
$N_3 = 10$ turns
$N_4 = 4$ turns
$N_5 = 18$ turns.

The actual numbers of turns used are determined for the most part by the size of the cores and by the desired signal and impedance levels. The signal voltage and signal current levels for this scale are about 5 volts and 0.5 amperes, respectively. The resistor value R is 5 ohms. The diode, D, must have a low forward impedance and a fast switching time although back resistance and recovery time are not critical.

The maximum speed of operation for the above-described circuits is largely determined by the characteristics of the core material and by the time required for resetting the input coupling cores. For this circuit, the speed is somewhat greater than 200 kc. The delay through the transfer stage is therefore about 5 μsec.

As expressed previously, the circuits are operated in a two-phase sequence. The clock cycle is divided into an A phase and a B phase. If the input to a connective occurs during the A clock phase, then the output occurs during the next B phase. Conversely, if the input occurs during the B clock phase, the output occurs during the next A phase. All circuit operation descriptions previously given have been for circuits whose input signals occur at SA time; output signals, if any, are

therefore present at *SB* time. For logical circuits designed to operate in the opposite phase, obvious permutations of drive pulses are necessary.

All circuits are packaged alike, and the manner in which the driver lines are connected determines the operating phase of a connective. As an example, consider a shift register which is constructed by connecting a number of transfer circuits in series. The current drivers are connected so that the first transfer output occurs at *A* time; the second, at *B* time; the third, at *A* time, etc. At *A* time all *A* phase circuits are read, and the information is transferred to the following *B* phase circuits. At *B* time all *B* phase circuits are read, and the information is transferred to the following *A* phase circuits. Information is alternately stored first in all the *A* phase connectives and then in all the *B* phase connectives. From this it should be recognized that information flows at a maximum rate of about 100 kc or at 10 μsec intervals.

Fig. 10 shows a hypothetical logical network such as might be employed for some machine control function. This particular configuration satisfies the expression

$$Y_{(n+4)} = P_{(n+2)} \cdot S_n \cdot \overline{T_{(n+1)}} \cdot \overline{U_{(n+1)}} \cdot \overline{V_{(n+1)}} \cdot (Q_n + R_n).$$

The subscripts n, $n+1$, etc., indicate the relative time of occurrence for each pulse. Fig. 10 is intended to illustrate the sequential nature of operations. For example, a *B* phase pulse on line *Q* at time (n) will next occur on line Y_1 at time $(n+1)$ which is an *A* phase pulse. In the absence of a concurrent pulse on line Y_2, a pulse will next appear on line Y_3 at time $(n+2)$. If a concurrent pulse appears on line *P*, a pulse will next appear on line Y_5 at time $(n+3)$. In the absence of a concurrent pulse on line Y_6, a pulse will appear on line Y_7 at time $(n+4)$.

V. Driver Design

The core circuits described are relatively insensitive to variations of *SA*, *SB*, *A* and *B* drive current pulses. Appreciable variations in current amplitude and waveform can occur without circuit failure. The circuits are most sensitive to the *RA*, *RB* and *DC* drivers, becoming more critical as the maximum speed of operation is approached. Accordingly, to obtain reasonable margins of operation, a slightly less than maximum operating speed is desirable. At 200 kc, amplitude variations of greater than ± 10 per cent are obtained for the most critical drivers.

The current driver described below performs satisfactorily with a minimum of regulation features. Adjustment of various parameters results in a driver with the following operating characteristics:

Rise time	0.1–0.5 μsec
Amplitude	1.0–2.2 ampere
Load characteristic for constant current operation	150.0 volts (back voltage).

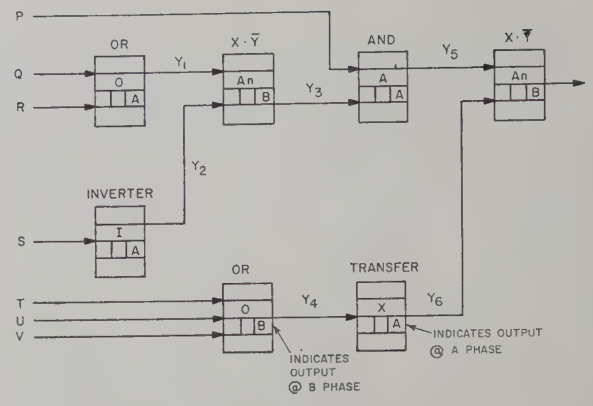


Fig. 10—Hypothetical logical network.

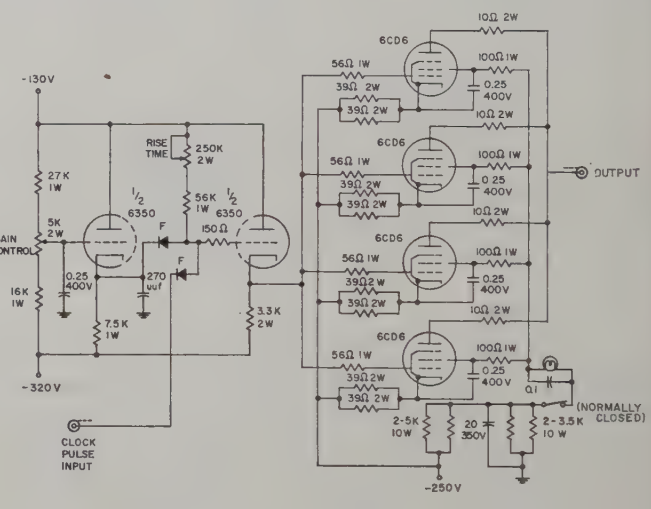


Fig. 11—Current driver.

The circuit is shown in Fig. 11. The core logic to be driven is connected in series with the common plate circuit of the current driver. The core circuits are returned to the zero volt bus, and this bus then provides the plate return voltage for the 6CD6 tubes.

Because the core logic circuits to be driven are series connected in the driver plate circuit, removal of a core card during testing will cause excessive screen current to flow. This condition occurs because the screen serves as a plate in the absence of plate potential. The lamp inserted in series with the screen supply then performs a dual purpose:

1) It limits screen current to safe value if plate potential disappears for any reason.
2) It provides an indication of an open plate circuit.

The switch shown in series with the screen supply is the on-off control for the driver package. During normal operation the switch is closed, connecting the screen supply to the common screen circuit.

The gain control sets the clipping level on the diode AND circuit by controlling the input to the first cathode follower. The rise time of the pulse is controlled by varying the diode AND circuit load resistance. The RC

of the circuit, formed by the input capacitance of the second cathode follower and the AND circuit load variable resistance, determines the rise time of the pulse but does not affect the fall time.

Cathode degeneration is used in the output stage for better regulation. When the back voltage builds up to the point where the plate current drops off, the screen conducts. The 0.25 μf capacitors from screen to cathode of the 6CD6's feed this negative voltage back to the cathode which serves to counteract the dropoff in plate current.

The *DC* current driver is obtained by applying a constant voltage grid drive to the 6CD6 tubes. Driver current data is described in Table I.

TABLE I
DRIVER CURRENT DATA

Driver	Nom. Rise Time μsec	Nom. Amplitude Amperes	Winding No. of Turns	MMF Ampere-Turns
A, B	0.35	1.8	3	5.4
SA, SB	0.35	1.8	5	9.0
RA, RB	0.5	1.25	2	2.5
DC	—	0.25	2	0.5

VI. Synthesized Circuit Functions

With the set of building blocks described in Section IV, it is possible to synthesize certain circuit functions necessary for a logical system. Some of these functions will now be described, using the symbolic representation of the building blocks.

A. OR Branch

This circuit is derived from the output branch circuit of one or more connectives. The logic is performed at the input to a connective and no additional machine time is required.

The circuits are connected in such a way that presence of a signal in either output (1) or (2) (see Fig. 12) will provide a ONE-input to the succeeding connective. Basically, this circuit is an "Inclusive OR" circuit. Isolation between the two output branch circuits is provided by the diode in each branch.

B. Dynamic Flip-Flop

Although the connectives have memory facilities for half a machine cycle, it is sometimes necessary in a machine to provide memory for an indefinite period of time. Therefore, a dynamic flip-flop is required. Fig. 13 shows three ways of obtaining the flip-flop operation.

In Fig. 13(a) an $X \cdot \overline{Y}$ and a transfer circuit are used. The output of the transfer is fed back to the input of the $X \cdot \overline{Y}$ in a closed-loop fashion. Fig. 13(b) shows a flip-flop using an $X \cdot Y$ and an OR connective. The operation is as mentioned before. Two $X \cdot \overline{Y}$ circuits can be used as shown in Fig. 13(c).

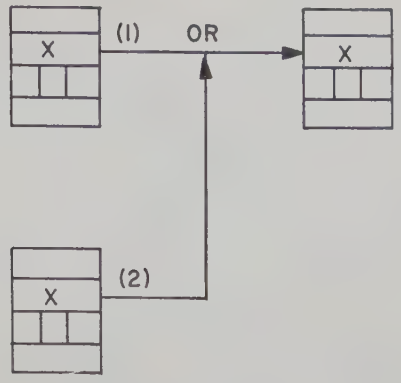


Fig. 12— OR branch.

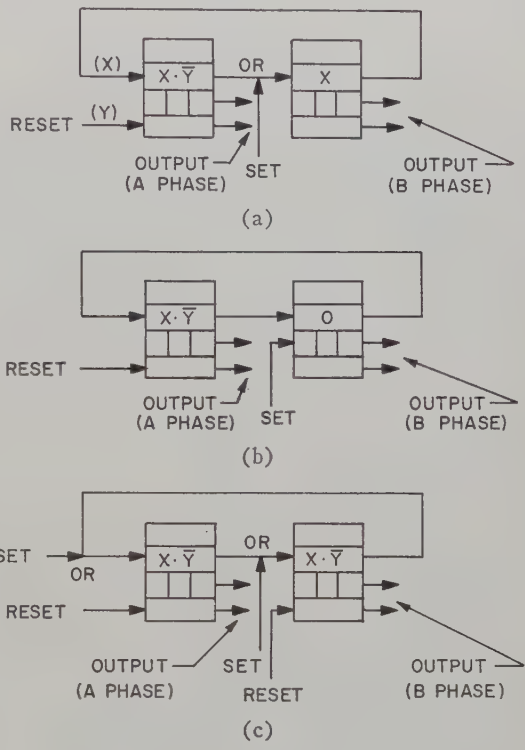


Fig. 13—Dynamic flip-flop.

C. Scale of Two Counter

This can be achieved by using two $X \cdot \overline{Y}$ circuits with inputs transposed and an OR circuit connected as shown in Fig. 14(a). An output pulse is obtained for every second input on output (1). Output (2) will have a pulse signal at every cycle time if an odd number of input pulses have been applied to the circuit.

D. Exclusive OR Circuit

The scale of two counter can be converted into an Exclusive OR circuit ($X \cdot \overline{Y} + \overline{X} \cdot Y$) if the feedback circuit is eliminated [see Fig. 14(b)]. The OR circuit can be replaced by an OR branch as previously described.

E. Delay Property

The basic transfer circuit has a delay equal to half a machine cycle. This property can be used in cases where

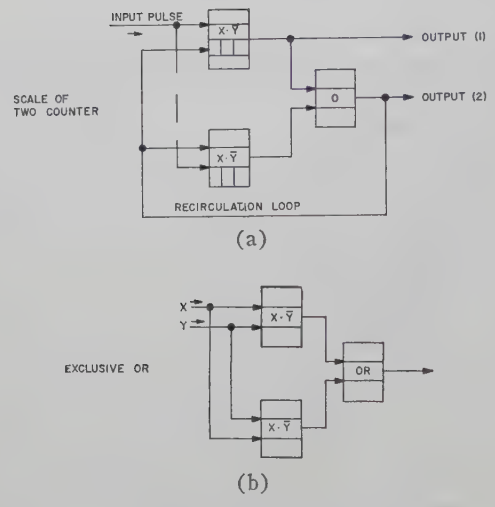


Fig. 14—Synthesized logical functions.

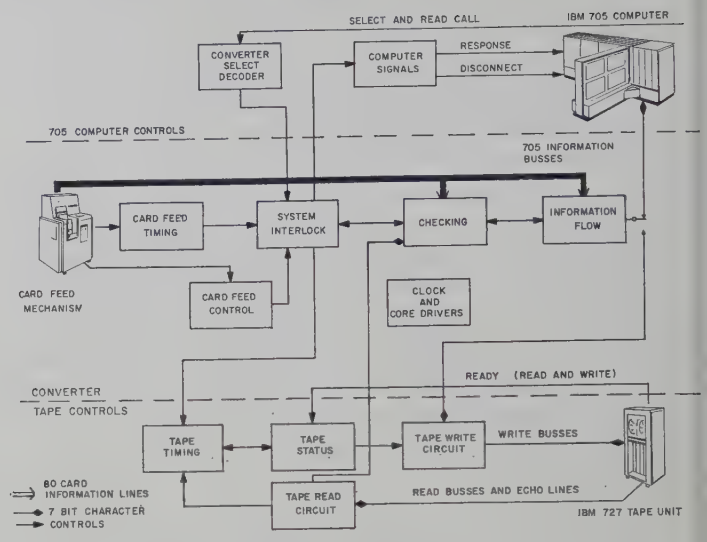


Fig. 15—714/759 card reader—synchronizer system.

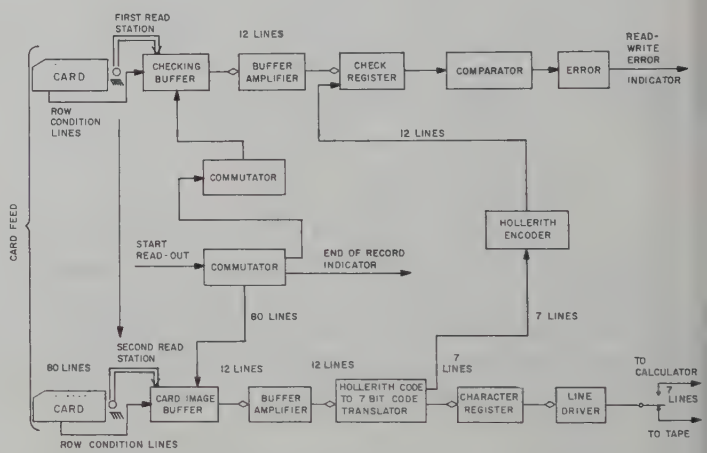


Fig. 16—Information flow.

time intervals must be established or where information storage is needed. In the latter case, a closed-loop shift register can be designed which circulates the information bits indefinitely. In the former case a closed ring circulating one index bit can be used, thus eliminating the need for single-shots or other delay devices.

VII. Data Flow Model System Design

Compared to synchronous vacuum tube-diode or transistor circuits, the core connectives under discussion exhibit special properties. The delay through a circuit is fixed and equal to half a machine cycle. Therefore, no delay lines or time standardizing circuits are needed. The clock and current drivers essentially perform the time standardization, and the basic transfer circuit can perform the function of one unit of delay. Furthermore, since each circuit has a gain of greater than unity, no signal-restoring circuits are needed.

To evaluate the circuits a data flow model was constructed and tested. A combined card-to-tape converter and card-to-computer input device, analogous in operation to the standard IBM 714/759 card synchronizer system, was chosen as the feasibility model. This particular system is well adapted to serve as a proving ground for new techniques because of its complexity. This complexity arises from the fact that three devices (the Tape Unit, the Card Feed, and the Electronic Computer) must be controlled and synchronized when operating.

A. Machine Organization

Fig. 15 shows the machine organization in functional form. The data-flow model can operate either with the 705 Computer or with the 727 Tape Unit. In either case it controls the card feed, the core buffer storage, the information flow, checking circuits, and the tape motion controls. The information flow is shown in Fig. 16.

Cards are read at the rate of 250 per minute. Each card is read twice and the card images are stored in two core buffers. The first time the card is read the complete card image is stored in a check buffer for checking purposes. The second time the card is read, the complete card image is stored in the information buffer. Card reading is done on a row-by-row basis which means that 80 brushes read the 80 possible holes in a card. If the machine is used as an input device to the 705 Computer, the information buffer is read out in a serial by character (column) fashion on demand from the program-controlled 705. The information is translated from the 12-bit code to a 7-bit 705 code. The character transfer rate is 33.6 μsec per character. After the buffer is emptied, *i.e.*, when all 80 characters are read out, a disconnect pulse is sent to the computer, signalling that the program can continue. The card reader is now activated and reads the next card into the buffer. If, at the end of the feeding operation, no Read Call from the computer is waiting, the machine latches up and waits for such a Read Call to occur.

The procedure for checking corrections of card reading is as follows. The card is read twice by two sets of brushes and the information is stored in two buffers. When the information is read out of the checking buffer, an odd-even bit count on a row-by-row basis is obtained and stored in a check register which consists of a set of 12 binary checking triggers. During reading of the information from the storage buffer, the characters are encoded back into 12-bit code and again an odd-even count is made. The two counts are compared at the end of the card cycle. If an unequal compare status is obtained, an error is signaled to the computer. Also, the transfer of information from the converter to the computer is checked in the computer by means of the 7-bit redundancy code.

For tape operation, the information flow and information reading from the card is the same as when reading into the 705. However, at the end of the card-read cycle the tape unit is started and the buffer contents are written onto a 727 tape. The character writing rate is 67 μsec per character. Checking is done in two ways: First, the double reading of the card checks for an odd-even count on a row-by-row basis, and then the echo from the tape write heads is checked for redundancy in the 7-bit configuration. When the record is stored on the tape, the tape unit is stopped and backspaced to the beginning of the record. Each character of the record is read and checked for redundancy. In addition, an odd-even count on a track-by-track basis is obtained for the record and compared to the odd-even count obtained previously. After this operation, the tape unit is stopped and the next card is read into the buffer. This will continue until the whole card file is stored on tape. During this time the card converter is in control of the tape unit; it senses the stopping and starting of the tape unit, the end of file condition on tape, etc.

All circuit functions indicated in the machine organization are performed by core logic circuits as described in the beginning of the paper. Of special interest is the commutator, which must perform two functions. First, it must remember which column has been read last and then it must furnish enough current to the buffer to read out the storage cores in each and every column. This is accomplished by a ring of transfer circuits. In series with the output core of the transfer circuit is the input core of the next transfer circuit and a particular drive line for a column of the buffer. As the transfer of a ONE from, say, stage (N) to stage ($N+1$) of the ring occurs, enough current in the output loop flows to reset the buffer cores. Therefore, no tubes or transistors or other active devices are used as current drivers for the core buffer. The checking register and comparator are built up of binary trigger circuits discussed previously.

The system discussed is highly reliable for various reasons. The cores are not subject to deterioration by outside influences nor do they undergo a change of characteristics with time. The additional components (one diode and one resistor), shown in the circuits, are used in a manner which is conducive to a long life span. The back voltage to which the diode is subjected, is well below the manufacturer's rating as is the forward current that flows through the diode. The back resistance of the diode is not critical and may vary from 1000 ohms to 1 megohm. Examination from a systems point of view indicates that the only parts of the machine which may deteriorate are the level converters and the current drivers themselves. However, these parts are easily maintained because only recurrent waveforms are observed for amplitude, rise time, and duration. For maintenance purposes, concern need not be given to the information status of the machine itself.

It is expected that this particular research study will open the way for incorporation of the described circuits into production machines. It is realized that technological improvements are possible in the driver circuits. At the present time vacuum tubes are used, but it is expected that in the near future these can be replaced by power transistors.

Since the data-flow model operation is analogous to the marketed IBM 714/719 employing vacuum tube-diode logic, a direct comparison of the component count can be made.

Components	*Tube-Diode Version*	*Core Version*
Tubes—Vacuum	475	200 (120 9 pin-miniatures) (80 octal base)
Tubes—Thyratron	100	10
Cams	50	7
Relays	40	14
Diodes	1190	600
Core Connectives	0	500

In the model the vacuum tubes, exclusive of the ones used to drive core logic circuits, perform level-changing functions. This is necessary because the core logic data flow model must communicate with a vacuum tube machine (705 Computer and 727 Tape Unit). For a self-contained, all-core logic system, level converters would not be necessary.

The above component count shows that a large reduction in cams and relays results. These electromechanical devices were used to perform circuit logic in the tube-diode version. For purposes of reliability these logic functions are now performed by core circuitry. The remaining cams are used to derive timing pulses to keep the mechanical feed in step with the electronic circuits.

VIII. Circuit Packaging and Data Flow Model Construction

The core circuits used in the data flow model are packaged on a specially designed card shown in Fig. 17. The card can contain up to six cores and will generally contain one logical connective. The card has 24 terminals which are used for all connections to each circuit.

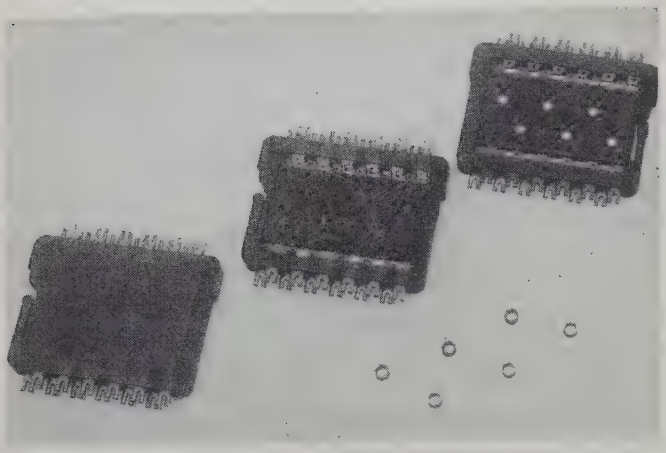


Fig. 17—Core circuit package.

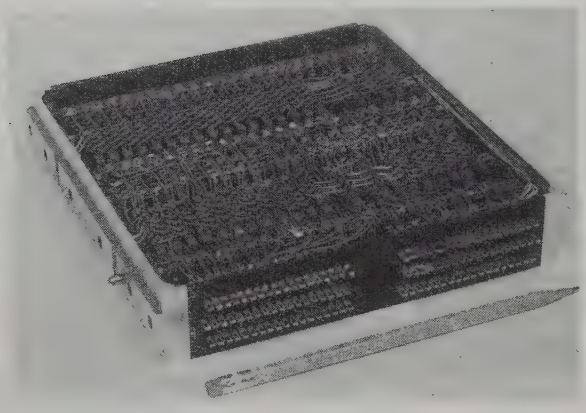


Fig. 18—Core card pluggable unit.

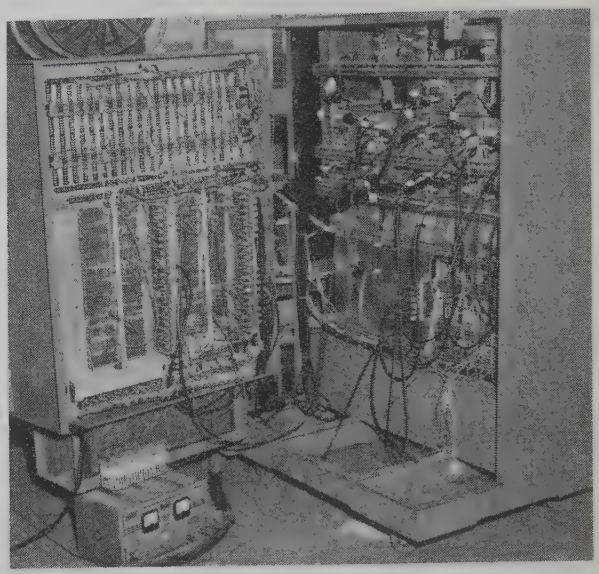


Fig. 19—Card-to-tape converter (pluggable unit gate from wiring side).

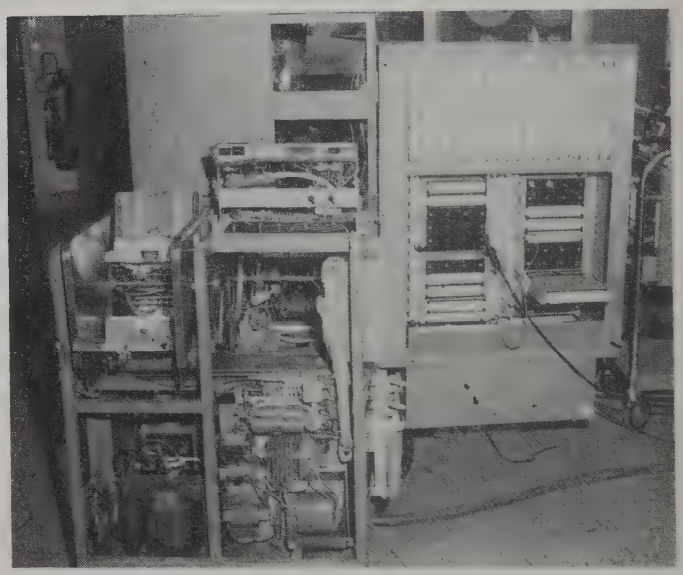


Fig. 20—Card-to-tape converter (pluggable unit gate from unit side).

For simplicity of wiring the same pair of terminals is assigned for each drive line on all cards. Where a drive line is not used, these terminals are shorted internally. Thus, fourteen terminals are reserved for drive currents *A*, *B*, *SA*, *SB*, *RA*, *RB* and *DC*, while six terminals are reserved for inputs, two for output, and two for the resistor. Where fewer than three inputs are used, the unused terminals are left blank.

Transfer circuits are wound two per card. This is practical for several reasons. The transfer circuit is limited to a single input and requires but three cores. Thus, there are sufficient terminals and core sockets available for two separate circuits. In addition, from a systems viewpoint a double transfer card can be used efficiently in ring, shift register, and in other applications where successive transfer operations occur. Two separate types of transfer cards were constructed, both of which were used to advantage in the machine. One card contains two transfers, both of which operate in the same phase. The other card contains two transfers which operate in alternate phases.

This construction lends itself to automatic assembly procedure and is also suitable for impregnation of the cores and windings. Impregnation with a thermosetting plastic insures winding protection from moisture, corrosive atmospheric conditions, and excessive handling. Since the thermoplastic has a dielectric constant of about four, it increases the breakdown voltage between windings over a value set by wire insulation alone.

When assembled and tested, the cards are inserted in a pluggable unit as shown in Fig. 18. The unit has provision for 48 circuit cards and their associated components (diode and resistor).

Figs. 19 and 20 show the data flow model assembly.

IX. Conclusions

The work described in this paper resulted in the design of a catalog of compatible logical building blocks suitable for use in the construction of a general logic system.

To evaluate the circuits a data flow model was built and tested. All circuit functions in the machine were performed whenever possible by core logic. Electro-mechanical devices, such as relays and cams, as well as vacuum tube and diode circuits, were replaced by toroidal core circuits. A substantial reduction in component count was obtained.

Principal advantages through the use of these circuits are:

1) Reliability: Ferrite cores do not change characteristics in normal usage or with age. This has already been proven by their large usage in computer memory applications.
2) Cost: Cores are inexpensive. Cost estimates, based on experience with the data flow model, indicate that low cost core logic circuitry is obtainable even when large quantity production is not realized.
3) Logical Design: Because the building blocks have built-in power gain, logical design is simplified. The circuits are synchronous, thus eliminating the need for reclocking or time restandardization. Building blocks of a relatively complex nature are available. This minimizes the circuits needed for a specified logical network. In addition, these circuits are compatible with core buffers commonly found in computers.
4) Serviceability: The data flow model proved that the testing of pulse type core logic systems does not present a problem. Sequential operation of circuits and series drive connections facilitate trouble shooting of the machine.

Limitations experienced in the design and use of these circuits are:

1) Power Supplies (Pulse Drivers): Suitable solid-state drivers were not readily available for usé in the data flow model. Power transistors of adequate power and speed were not available. New commercial transistors recently developed show, however, potential application for power drivers.
2) Packaging: The core cards used were hand wound. Automation of this operation is proposed, which will result in considerable savings. In addition, further work is required with regard to methods of potting.
3) Temperature Sensitivity: In common with other solid-state devices, these circuits require that the temperature be held in a controlled range. The problem is simplified, however, because the controlled range need not be centered about room temperature. In the model, small fans insured movement of air past the cores with satisfactory results.
4) Speed: The maximum speed is necessarily a compromise of several factors, including size, drive power, and operating impedance and voltage levels. No great effort was made for this model to increase the speed of circuit operation.

For accounting machines, input/output devices, and small calculators, the speed of toroidal core logic is sufficient. Together with their inherent reliability and inexpensive production costs, they will prove useful in the future.

X. Acknowledgment

Contributions to the work reported in this paper were made by L. G. Allen, E. A. Daley, and G. L. Hotchkiss in the systems design, A. A. Kahn in the circuit design, and H. J. Hallstead in packaging. Their efforts are sincerely acknowledged.

The Use of a Repetitive Differential Analyzer for Finding Roots of Polynomial Equations*

P. MADICH†, J. PETRICH†, AND N. PAREZANOVICH†

Summary—**The paper describes a procedure for obtaining real and complex roots of algebraic equations with real or complex coefficients by the use of a repetitive differential analyzer. The procedure requires only operational amplifiers and ganged linear potentiometers. Differential analyzers are very suitable for solving algebraic equations since they permit visual checking of the procedure and make it possible to investigate how the roots of the polynomial are affected by variation of its coefficients. The procedure is not iterative.**

Introduction

THE solution of algebraic equations of higher degree (frequently encountered as a practical problem) has received considerable attention. Many numerical and graphical methods as well as a number of special machines[1–3] are available for solving this problem. For this purpose it is also possible to use machines which are intended for the solution of other problems.[4,5] Differential analyzers are also used for obtaining the roots of polynomials. The efficiency of using such machines depends on the procedure applied. The simplest case is that in which only real roots of the polynomial are required. By some procedures[5] it is then possible to obtain the solutions directly while other procedures require certain adjustments. However, the solution is complicated whenever complex roots are involved, because in the majority of cases it is then necessary to apply iterative procedures. Special difficulties are encountered when polynomials with complex coefficients are to be solved. Equipment for solving such a problem[4] requires a large number of elements ($5n$ operational amplifiers) and the procedure is iterative. Certain machines also enable multiple roots to be detected, but this problem has not yet been solved in a general way.

The procedure described in this paper uses a repetitive differential analyzer for finding the roots of polynomials with either real or complex coefficients and makes it possible for both real and complex roots, single as well as multiple, to be obtained. Accordingly, this procedure embraces all cases encountered in solving for the roots of polynomial equations. Because the procedure is not iterative, the repetitive analyzer, thanks to its well-known advantage of providing direct and continuous display of solutions, makes it possible to vary individual parameters and examine their effect on the roots of a polynomial. The accuracy of results is within the limits obtainable in solving standard problems and is directly dependent on the accuracy of operational amplifiers, ganged potentiometers and the measuring system. The accuracy can further be increased by the use of suitable transformations. In general, analyzing the existing procedures which use various elements of analog computing technique as auxiliary means, it is seen that the repetitive differential analyzer provides wide possibilities of solving for the roots of polynomials with a simple procedure.

Mathematical Bases of the Procedure

The most general case of solving for the roots of polynomials is the one in which both the coefficients and roots are complex. Such cases are the most difficult ones but, fortunately, they are seldom encountered in practice. In the majority of cases it is necessary to deal with the special case of polynomials having real coefficients, and both real and complex roots. Finally, the extra special case is represented by polynomials with coefficients and roots that are both real.

Polynomials with Real Coefficients and Roots

The general form of a polynomial with real coefficients and roots is

$$f(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n. \qquad (1)$$

By differentiating (1) n times with respect to x the differential equation of nth order is obtained.

$$f^{(n)}(x) = n! \cdot a_n \qquad (2)$$

with the following initial conditions for $x = 0$.

$$f^{(k)}(0) = k! \cdot a_n \qquad k = 0, 1, 2, \cdots, n-1. \qquad (3)$$

The setting up of differential equation (2), on a differential analyzer is a simple matter and the solution obtained is a parabola described analytically by (1). The roots of polynomial (1) are obtained directly as the intersections of the parabola with the x-axis. This procedure is used with both repetitive and dc differential analyzers.[5]

To determine multiple roots, use is made of the well-known conditions which the derivatives of the function must satisfy at the point representing a multiple root. The nature of the root, whether it is single or multiple, is indicated by the intersection of the parabola with the x-axis. Thus, for example, in the case of a double root,

* Manuscript received by the PGEC, October 7, 1958.
† Inst. of Nuclear Sciences, Boris Kidrich, Belgrade, Yugoslavia.
[1] A. S. Householder, "Principles of Numerical Analysis," McGraw-Hill Book Co., Inc., New York, N. Y.; 1953.
[2] J. S. Frame, "Machines for solving algebraic equations," *MTAC*, vol. 1, pp. 337–353; 1954.
[3] N. N. Mikhaelov (Russian), "Electrical devices for solving algebraic equations," *Automatika i Telemekanika*, vol. 19, pp. 477–490; 1958.
[4] F. W. Bubb, Jr., "A circuit for generating polynomials and finding their zeros," Proc. IRE, vol. 39, pp. 1556–1561; December, 1951.
[5] C. Atkinson, "Polynomial Root Solving on the Electronic Differential Analyser," *MTAC*, vol. 9, pp. 139–142; 1955.

the parabola only touches the x-axis at this point and has either a maximum or a minimum which is clearly indicated on the differential analyzer. By examining the first derivative of the polynomial, which is easily obtained on a repetitive differential analyzer, it can be established whether the examined point also represents a root of the first derivative of the polynomial. The procedure for examining for roots of higher multiplicity is evident.

Polynomials with Real Coefficients and Real and Complex Roots

The general form of this polynomial is

$$f(z) = a_0 + a_1 \cdot z + a_2 \cdot z^2 + \cdots + a_n \cdot z^n \tag{4}$$

where $a_j (j=0, 1, 2, \cdots, n)$ are real numbers and $z=x+i\cdot y$. Expression (4) can be expanded in Taylor's series:[1]

$$f(z) = f(x) + i\cdot\frac{y}{1!}\cdot f'(x) - \frac{y^2}{2!}\cdot f''(x) + \cdots + i^n\cdot\frac{y^n}{n!}\cdot f^{(n)}(x). \tag{5}$$

Grouping together real and imaginary terms in (5) one obtains

$$f(z) = u(x, y) + i\cdot v(x, y) \tag{6}$$

where

$$u(x, y) = f(x) = \frac{y^2}{2!}\cdot f''(x) + \frac{y^4}{4!}\cdot f^{(4)}(x) - \frac{y^6}{6!}\cdot f^{(6)}(x) + \cdots \tag{7}$$

$$v(x, y) = y\cdot f'(x) - \frac{y^3}{3!}\cdot f'''(x) + \frac{y^5}{5!}\cdot f^{(5)}(x) - \cdots \tag{8}$$

The values of x and y that simultaneously satisfy the conditions

$$u(x, y) = 0 \tag{9}$$

$$v(x, y) = 0 \tag{10}$$

represent the real and the imaginary part of a complex root of the polynomial (4) respectively. Expressions (7) and (8) have already been used in solving for the roots of polynomials with the aid of a differential analyzer.[5] They can be written in a more convenient form

$$u(x, y) = f(x) - y^2\left\{\frac{f''(x)}{2!} - y^2\left[\frac{f^{(4)}(x)}{4!} - y^2\left(\frac{f^{(6)}(x)}{6!} - \cdots\right)\right]\right\} \tag{11}$$

$$\frac{v(x, y)}{y} = f'(x) - y^2\left\{\frac{f'''(x)}{3!} - y^2\left[\frac{f^{(5)}(x)}{5!} - y^2\left(\frac{f^{(7)}(x)}{7!} - \cdots\right)\right]\right\}. \tag{12}$$

Expressions (11) and (12) are very suitable for presentation on a differential analyzer as will be explained later.

If a pair of conjugate complex roots of a polynomial is to be multiple, where, for example, $z_r = x_r \pm i\cdot y_r$ is a multiple root of kth order (k being less than $n/2$), then it is necessary and sufficient that the following conditions be fulfilled

$$\frac{\partial^p u(x, y_r)}{(\partial x)^p} = 0 \qquad p = 0, 1, 2, \cdots, k-1 \tag{13}$$

$$\frac{\partial^p v(x, y_r)}{y_r(\partial x)^p} = 0 \qquad p = 0, 1, 2, \cdots, k-1 \tag{14}$$

for $x=x_r$. From the graphs of (11) and (12) it is possible to determine on the basis of their intersections with the x-axis whether they represent single or multiple roots. By further investigating the partial derivatives of polynomials (11) and (12) with respect to x for $x=x_r$ and $y^2=y_r^2$, it is possible to determine the multiplicity of the complex root.

Polynomials with Complex Coefficients and Roots

Their general form is

$$f(z) = c_0 + c_1\cdot z + c_2\cdot z^2 + \cdots + c_n\cdot z^n. \tag{15}$$

where $c_j (j=0, 1, 2, \cdots, n)$ and z are complex numbers. Substituting $c_j = a_j + i\cdot b_j$ and grouping together real and imaginary terms one obtains

$$f(z) = f_a(z) + i\cdot f_b(z) \tag{16}$$

where

$$f_a(z) = a_0 + a_1\cdot z + a_2\cdot z^2 + \cdots + a_n\cdot z^n \tag{17}$$

$$f_b(z) = b_0 + b_1\cdot z + b_2\cdot z^2 + \cdots + b_n\cdot z^n. \tag{18}$$

Each of the polynomials (17) and (18) can be represented by two polynomials in the same way as (4) has been represented by (7) and (8). One has then

$$f_a(z) = u_a(x, y) + i\cdot v_a(x, y) \tag{19}$$

where

$$u_a(x, y) = f_a(x) - \frac{y^2}{2!}\cdot f_a''(x) + \frac{y^4}{4!}\cdot f_a'''(x) - \cdots \tag{20}$$

$$v_a(x, y) = y\cdot f_a'(x) - \frac{y^3}{3!}\cdot f_a'''(x) + \frac{y^5}{5!}\cdot f_a^{(5)}(x) - \cdots \tag{21}$$

and

$$f_b(z) = u_b(x, y) + i\cdot v_b(x, y) \tag{22}$$

where

$$u_b(x, y) = f_b(x) - \frac{y^2}{2!}\cdot f_b''(x) + \frac{y^4}{4!}\cdot f_b^{(4)}(x) - \cdots \tag{23}$$

$$v_b(x, y) = y\cdot f_b'(x) - \frac{y^3}{3!}\cdot f_b'''(x) + \frac{y^5}{5!}\cdot f_b^{(5)}(x) - \cdots. \tag{24}$$

Substituting (19) and (22) in (16) one obtains

$$f(z) = [u_a(x, y) - v_b(x, y)] + i[v_a(x, y) + u_b(x, y)]. \tag{25}$$

From (25) taking into account (20), (21), (23) and (24) it follows that

$$u(x, y) = f_a(x) - y \cdot f_b'(x) - \frac{y^2}{2!} \cdot f_a''(x) + \frac{y^3}{3!} \cdot f_b'''(x) + \frac{y^4}{4!} \cdot f_a^{(4)}(x) - \cdots \quad (26)$$

$$v(x, y) = f_b(x) + y \cdot f_a'(x) - \frac{y^2}{2!} \cdot f_b''(x) - \frac{y^3}{3!} \cdot f_a'''(x) + \frac{y^4}{4!} \cdot f_b^{(4)}(x) + \cdots . \quad (27)$$

It is more convenient for setting on a differential analyzer if the polynomials $u(x, y)$ and $v(x, y)$ are in the following form,

$$u(x, y) = f_a(x) - y\left\{f_b'(x) + y\left[\frac{f_a''(x)}{2!} - y\left(\frac{f_b'''(x)}{3!} - \cdots\right)\right]\right\} \quad (28)$$

$$v(x, y) = f_b(x) + y\left\{f_a'(x) - y\left[\frac{f_b''(x)}{2!} + y\left(\frac{f_b'''(x)}{3!} - \cdots\right)\right]\right\}. \quad (29)$$

The real and the imaginary part of the complex root of polynomial (15) are obtained from the conditions

$$u(x, y) = 0 \quad (30)$$

$$v(x, y) = 0. \quad (31)$$

Multiple roots of polynomials with complex coefficients are obtained in a similar manner, as in the case of real coefficients. Namely, if $z_r = x_r + i \cdot y_r$ is a multiple root of kth order ($k < n$), then it is necessary and sufficient if the following conditions are satisfied

$$\frac{\partial^p u(x, y_r)}{(\partial x)^p} = 0 \qquad p = 0, 1, 2, \cdots, k-1 \quad (32)$$

$$\frac{\partial^p v(x, y_r)}{(\partial x)^p} = 0 \qquad p = 0, 1, 2, \cdots, k-1 \quad (33)$$

for $x = x_r$.

Use of the Repetitive Differential Analyzer

Polynomials with Real Coefficients and Roots

The roots of polynomials with real coefficients and roots (1) are obtained directly from the intersections of the x-axis and the parabola representing the solution of the differential equation (2), with the initial conditions (3). This procedure has already been applied using differential analyzers. A block diagram for solving (2) is shown in Fig. 1. With such a circuit one obtains not only the parabola but also the derivatives of the polynomial, which serve to determine multiple roots.

Polynomials with Real Coefficients and Real and Complex Roots

In solving this kind of problem on a repetitive analyzer, use is made of (11) and (12) which correspond to polynomial (4). Polynomial $f(x)$ and its derivatives are obtained from the circuit of Fig. 1. The magnitude y^2 is considered as a variable parameter t and is represented by means of a linear potentiometer. The solution of such a problem on the differential analyzer requires n summing amplifiers and $n-1$ ganged linear potentiometers in addition to the n integrators necessary for generating parabola $f(x)$ and its derivatives. The block diagram is shown in Fig. 2. Adjustment of the parameter t to a value for which parabolas (11) and (12) have common intersections with the x-axis can be done directly by observing the two parabolas simultaneously on a double beam oscilloscope. No iterative procedure is required.

To investigate multiple roots on the differential analyzer it is necessary to use a different circuit giving partial derivatives of polynomials (11) and (12) with respect to x. Such a circuit obviously requires less elements than the circuit of Fig. 2.

Polynomials with Complex Coefficients and Roots

To solve this kind of problem on a repetitive differential analyzer, use is made of (28) and (29) which correspond to the polynomial (15). In order to generate polynomials $f_a(x)$ and $f_b(x)$, defined by (17) and (18), a total of $2n$ integrators is required. The circuit simulating (28) and (29) requires $2n$ summing amplifiers and slightly more than $2n$ ganged linear potentiometers. The block diagram is shown in Fig. 3. (Note that the multiplication of each term $f^{(i)}(x)$ by $1/i!$ has not been explicitly indicated.) Adjustment of the parameters can be done directly, observing simultaneously parabolas (28) and (29).

If multiple roots are to be investigated, it is easy to construct circuits simulating the derivatives of polynomials (28) and (29).

In some cases the differential analyzer requires certain suitable transformations to be made in order to make sure the investigated problem will not exceed the working interval of the computer. Transformations that are necessary in solving for the roots of polynomials on a differential analyzer have already been treated.[4]

Examples

In order to illustrate the procedure, three examples, one for each of the three kinds of problems, have been worked out. They were solved on the repetitive differential analyzer.[6] Measurements were made on a special measuring system,[7] while a cathode-ray oscilloscope served for visual observation of the procedure.

Example 1

Solve the roots of the polynomial

$$f(x) = x^3 - 1.9 \cdot x^2 + 0.99 \cdot x - 0.081 = 0.$$

[6] R. Tomovich, and D. Mitrovich, "Some experiences with a repetitive differential analyser," *Bull. Inst. Nuclear Sci., Boris Kidrich*, Belgrade, Yugoslavia, vol. 8, pp. 109–116; 1958.

[7] T. Aleksich, "Measuring of instantaneous values of periodic voltage wave forms," *Bull. Inst. Nuclear Sci., Boris Kidrich*, Belgrade, Yugoslavia, vol. 3, pp. 127–130; 1953.

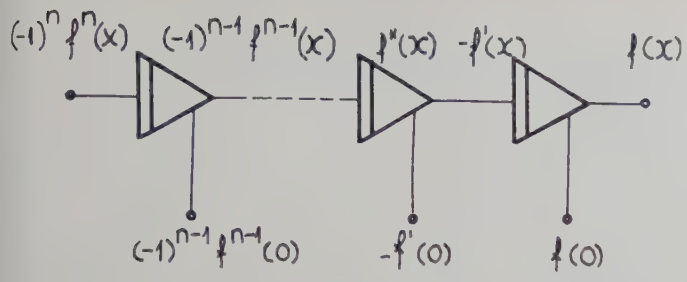


Fig. 1—Block diagram for solving equations with real coefficients and roots.

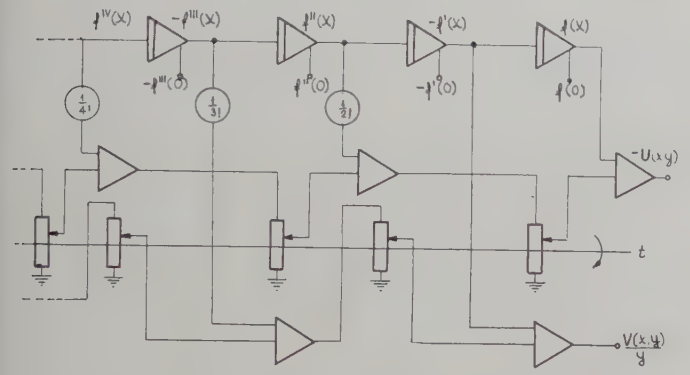


Fig. 2—Block diagram for solving equations with real coefficients and complex roots.

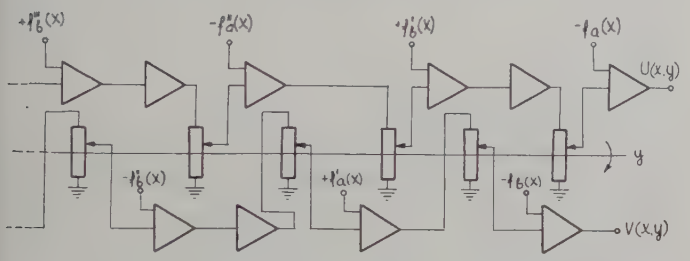


Fig. 3—Block diagram for solving equations with complex coefficients and roots.

This function is determined by the third-order differential equation

$$f'''(x) = 6$$

and by the initial conditions $f(0) = -0.081$; $f'(0) = +0.99$; $f''(0) = -3.8$. The following results were obtained on the analyzer. $x_1 = +0.096$; $x_2 = x_3 = +0.907$ while the exact values are $x_1 = +0.1$; $x_2 = x_3 = +0.9$.

The results as obtained on the cathode-ray oscilloscope are shown in Fig. 4.

Example 2

Find the roots of the polynomial

$$f(z) = z^5 - 9.02 \cdot z^3 + 21.96 \cdot z^2 - 21.6199 \cdot z + 8.8804 = 0.$$

This polynomial is determined by the differential equation

$$f^{(5)}(z) = 120$$

and the initial conditions $f(0) = +8.8804$; $f'(0) = -21.6199$; $f''(0) = +43.92$; $f'''(0) = -54.12$; $f^{(4)}(0) = 0$. By solving this problem on the differential

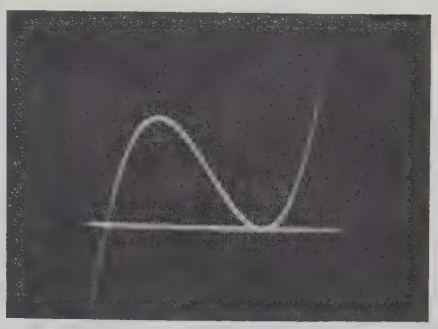


Fig. 4—Solution of equation $x^3 - 1.9 \cdot x^2 + 0.99 \cdot x - 0.081 = 0$ as obtained on a cathode-ray oscilloscope.

analyzer a double pair of conjugated complex roots $z_{1,2} = z_{3,4} = 0.99 \pm i \cdot 0.72$ and a real root $z_5 = -4.03$ are obtained. The exact values are $z_{1,2} = z_{3,4} = 1 \pm i \cdot 0.7$; $z_5 = -4$.

Example 3

Find the roots of the polynomial with complex coefficients:

$$f(z) = z^2 - z(1.8 + i \cdot 0.8) + (0.68 + i \cdot 0.76) = 0.$$

After separating real and imaginary parts, two polynomials with real coefficients are obtained

$$f_a(z) = z^2 - 1.8z + 0.68$$
$$f_b(z) = -0.8z + 0.76.$$

These polynomials are simulated by solving the differential equation

$$f_a''(z) = 2 \qquad f_a(0) = 0.68;\ f_a'(0) = -1.8.$$
$$f_b'(z) = -0.8 \qquad f_b(0) = 0.76.$$

The following results were obtained on the analyzer: $z_1 = 1.01 + i \cdot 0.19$; $z_2 = 0.807 + i \cdot 0.603$. The exact values are $z_1 = 1 + i \cdot 0.2$; $z_2 = 0.8 + i \cdot 0.6$.

Conclusion

Repetitive differential analyzers have proved to be very useful for several reasons:

1) The setting up of problems is straightforward and easy.

2) Only standard operational amplifiers and other linear elements are used, no electronic multipliers and function generators being required. It is thus possible with a minimum of equipment to find all roots of a polynomial, up to the 10th order, with good accuracy.

3) Advantages of this procedure are obvious since it enables solution of the generalized case without limitations concerning coefficients and roots.

4) The effect of various parameters on the solutions is easily analyzed.

5) The procedure is not iterative.

Acknowledgment

The authors wish to express their thanks to Dr. Rajko Tomovich for his interest in this work and for valuable suggestions.

A High-Speed Analog-Digital Computer for Simulation*

R. C. LEE† AND F. B. COX†

Summary—**This paper describes the principles of operation and logical design of an analog-digital computer capable of simulating complex physical systems in real time. Information in the machine is represented by an analog voltage pulse and a digital number. Arithmetic operations are performed in time-shared analog computing components and conventional digital logical elements. A novel floating-point arithmetic feature is provided to extend the dynamic range of the machine variables.**

Instructions and constants are stored on a magnetic drum before computation begins. The instructions determine the sequence of computer operations, and both the instructions and constants are arranged so that random access to the drum is not needed.

The programming techniques developed for the computer are described. The inherent simplicity of these techniques should permit engineers directly concerned with simulation to program their own problems for computer solution.

Introduction

THIS paper describes an analog-digital computer which was designed for the real-time simulation of physical systems. It is, therefore, especially suited for the high-speed solution of complex sets of nonlinear differential equations. Problem variables are represented in the computer in combined analog and digital form, and high-speed, time-shared analog computing elements are used to save equipment. In order to extend the computer's dynamic range, to improve accuracy, and to provide ease of programming, a novel floating-point arithmetic feature is also included in the design. Organization of the computer is such that the programming techniques are easily learned and applied.

A simplified prototype of this computer has been constructed to evaluate the basic principles of this form of computation, *i.e.*, the use of time-shared analog computing elements for performing arithmetic operations in the computer. Despite the fact that the floating-point feature has not yet been installed, solution accuracies for a second-order differential equation solved in a computation interval of 200 μsec ranged from one to three per cent.

Because of its analog nature, the computer's arithmetic element is easily accessible to input and output signals in the analog domain. The simplicity achieved through the use of time-shared analog computing components makes it economically feasible to effectively increase the computing speed by operating two or more arithmetic elements in parallel. Although simpler in construction, the arithmetic element cannot achieve the precision that would be obtainable with a digital arithmetic element. The use of analog computing elements in the computer should result in over-all computational accuracies of from one to five per cent.

The computer to be described, which will be referred to as the analog-digital computer, represents a somewhat novel departure from existing and proposed simulation computers. At the present time, most simulation problems—and especially situations in which computations must keep pace with real time—are run on analog machines.[1] Analog computers, however, require a large number of computing elements and are limited in both dynamic range and accuracy. Considerable effort, therefore, is being spent currently to develop simulators which employ either digital or combined analog-digital techniques.[2,3]

Mixed analog-digital systems generally take the form of two separate machines with conversions between the two made at appropriate points. The usual procedure is then to perform part of the computations in the digital machine and part of the computations in the analog machine. The analog-digital computer, however, is a single, integrated system, and computations are performed in a combined analog-digital domain.

The principles of operation and the basic design of the analog-digital computer are presented in this paper.[4] Also, the general nature of the programming techniques developed for the computer are described. No knowledge of the computer's logic or the detailed techniques for programming a general-purpose digital computer is required. Finally, the analog-digital computer is evaluated in terms of the class of problems that can be solved in real time, the complexity of the system, and the effort required to prepare a problem for computer solution.

* Manuscript received by the PGEC, October 6, 1958; revised manuscript received, February 18, 1959. This paper is based in part upon a thesis submitted by R. C. Lee in May, 1958, to the Department of Electrical Engineering, Massachusetts Institute of Technology, in partial fulfillment of the requirements for the degree of Master of Science, and upon research supported by the Navy under contract with M.I.T.

† Servomechanisms Lab., Dept. of Elect. Engrg., M.I.T., Cambridge, Mass.

[1] "Real time" in this paper is intended to imply high-speed computation. As an example, a computation cycle of 20,000 μsec is required to provide accurate and stable solutions for the real-time simulation of a high-performance aircraft.

[2] W. H. Dunn, C. Eldert, and P. V. Levonian, "A digital computer for use in an operational flight trainer," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-4, pp. 55–63; June, 1955.

[3] R. M. Leger and J. L. Greenstein, "Simulate digitally, or by combining analog and digital computing facilities," *Control Engineering*, vol. 3, pp. 145–153; September, 1956.

[4] For a detailed description of the logical design, see R. C. Lee, "Logical design of a high-speed analog-digital computer for simulation," Master's thesis, Dept. Elect. Engrg., M.I.T.; May, 1958.

System Organization

In the analog-digital computer, physical quantities are represented in part by the amplitude of an analog voltage and in part by a digital number. Analog voltages are not continuous time functions but are pulses whose amplitudes are equal to the scaled magnitude of the quantities represented. Digital numbers in the computer represent the scale factors associated with the machine variables. This is illustrated by the scaling equation

$$Y = Ay$$

where

Y = true magnitude
y = scaled magnitude expressed in machine units
A = scale factor.

The scale factor of y is chosen to be

$$A \geq \text{maximum expected value of } |Y|$$

since in machine units

$$y \leq 1.0.$$

In the analog-digital computer, scaling is done in powers of ten, *i.e.*,

$$A = 10^x.$$

The amplitude-modulated analog pulse represents y machine units, and the digital number is the binary representation of the exponent x. Analog voltages in the machine are always positive, and therefore a sign bit is included in the digital word.

The computer system block diagram is shown in Fig. 1. In the multiplexed analog arithmetic element, analog computing components are time-shared, and operations on the machine variables are performed in time sequence. The digital system simultaneously performs operations in a digital accumulator on the scale factors of the machine variables. The sequence of operations is determined by a program stored in the magnetic drum storage element. Operations performed by both the analog and the digital systems are directly controlled by the digital control element. Analog inputs and outputs are read directly into and out of the analog arithmetic element.

As shown in Fig. 1, the program and data may be read into the computer prior to the start of computation either automatically from a tape reader or manually from a keyboard. When the switches shown are in the COMPUTE position, however, the writing circuits are disconnected, and no further information is written on the drum. Because no information may be written on the drum during computation, a small capacity (70 bits) magnetic core memory is used in the digital system to store the scale factor and sign of intermediate results obtained during computation.

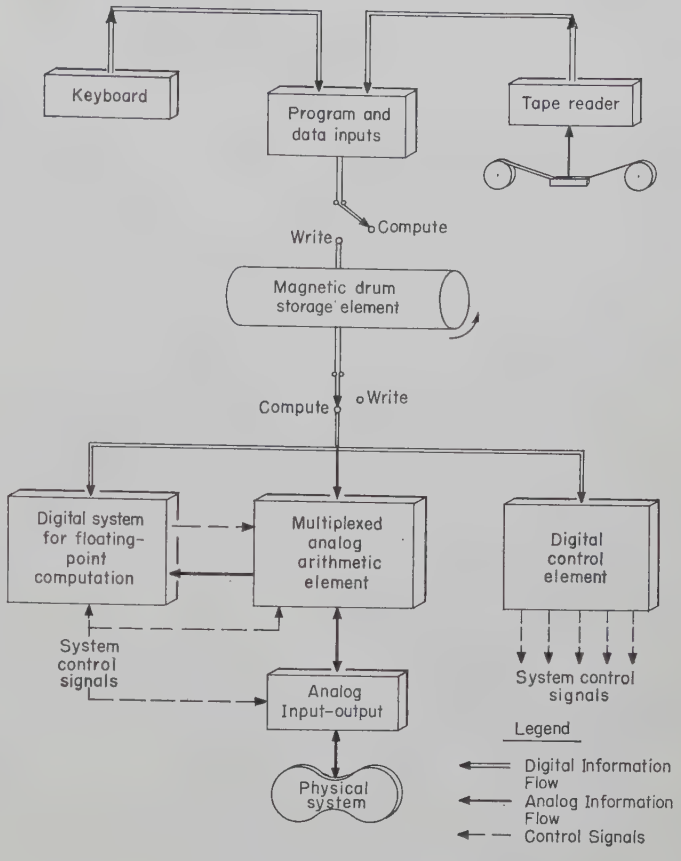


Fig. 1—System block diagram.

The program and data are stored on the drum in the order required by the computer, and so there is no need for random access to the drum. Data words transferred to the analog arithmetic element are first decoded into an analog voltage by a digital-to-analog decoder. Data words sent to the digital system or instructions sent to the digital control element, however, are transferred directly in digital form.

In order to extend the dynamic range of the computer and to simplify programming procedures, a floating-point arithmetic feature is included to automatically scale the machine variables into an optimum range during computation. Because scaling is done in powers of ten, the optimum range of analog voltages has been defined as one-tenth to one machine unit. Analog voltages are adjusted to lie within this optimum range of values, and the digital exponents are correspondingly modified in the following manner: 1) The present level of the analog voltage is determined, and information describing this level is transferred from the analog arithmetic element to the digital system. 2) Control signals are sent from the digital system to multiply the analog voltage by the power of ten necessary to place it within the optimum range. 3) The digital exponent is corrected by the accumulator in the digital system. This method of automatic scale-factor correction makes it possible to handle optimum quantities in the machine and there-

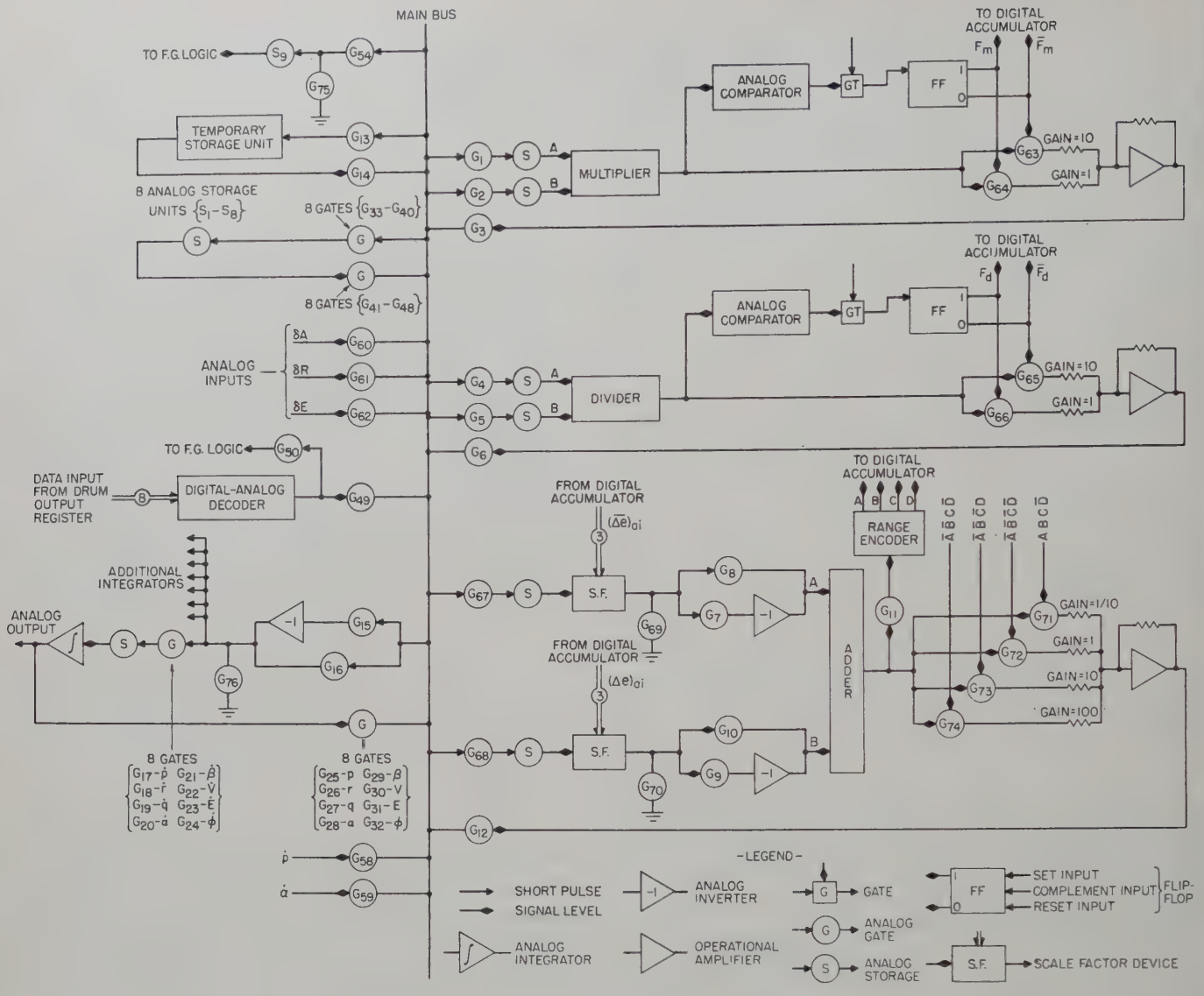


Fig. 2—Multiplexed analog arithmetic element.

by overcomes a major limitation of analog computers, namely a limited dynamic range.[5]

Multiplexed Analog Arithmetic Element

The multiplexed analog arithmetic element, shown in Fig. 2, performs all arithmetic operations on the analog machine variables. Single units are provided to multiply, to divide, and to add (or subtract). The temporary storage element is provided to store intermediate results, and the analog storage units (S_1 through S_8) are used both for the generation of arbitrary functions (see Appendix II) and as additional temporary storage elements. Analog integrators are indicated in the figure since they are presently being employed. Other integration methods may be used, however, as discussed in Appendix I. Because analog integrators cannot be time-shared, a separate unit must be provided for each computed derivative.

Information in the analog arithmetic element is transferred over a single bus as a short-duration voltage pulse. The clock frequency of the computer is 100 kc, and therefore the voltage pulse is somewhat less than 10 μsec in duration. For each transfer operation, the digital control element enables two analog gates—one at the origin of the data and the other at the destination. Storage units are provided at the input to the computing elements, such as the multiplier, to temporarily store one input while the second analog voltage is being transferred to the other input.

Analog inputs to the computer (δA, δR, and δE) are gated directly onto the main bus. These inputs, which arise from external sources, may be sampled when needed in a manner identical to the sampling of internal

[5] It should be pointed out here that analog input variables have a fixed scale factor, and samples of these voltages, when read into the computer, will generally not be in the defined optimum range. Samples of the input variables therefore are optimized by the procedure just described before being used in the computation. This procedure is discussed in more detail in the next section.

data. As was pointed out in the previous section, these analog inputs have fixed scale factors which are chosen to prevent their maximum voltage amplitudes from exceeding one machine unit and which are preset in the magnetic core memory. Because an input sample may lie anywhere between zero and one machine unit, it is necessary to adjust the sampled voltage to lie in the optimum range of one-tenth to one machine unit after it is read into the machine. This optimization procedure is described later in this section. Output information is, in most cases, the result of an integration. Hence, if analog integrators are used as indicated in Fig. 2, the computer outputs may be read out as continuously varying analog voltages having fixed scale factors.

Fifteen ranges of analog voltages, corresponding to exponents defined by four binary bits, are used for scaling in the computer. The range of numbers that may be handled in the computer is 10^{-7} to 10^{7}. Hence, in the digital word, three bits represent the magnitude and one bit represents the sign of the exponent. Numbers less than 10^{-14} machine unit are regarded as zero by the computer and are designated as such by a unique digital number, referred to as the zero-code.

Information which is stored in digital form on the magnetic drum is read into a digital-to-analog decoder. The decoder output is an analog voltage which may be gated onto the main bus through G_{49}, as indicated in Fig. 2. The method of data storage on the drum insures that the analog voltage will be available at the decoder output when it is needed by the computer.

The analog multiplier is provided for multiplication of two analog voltages. If it is desired, for example, to form the product xy in the computer, the analog quantity x is sampled from a point in the machine by enabling a particular gate, and the voltage pulse is transferred to the A-input of the multiplier by enabling G_1 (Fig. 2). Hence, the quantity x is stored in an analog storage device. Simultaneously, the digital word associated with x (*i.e.*, the digital exponent x_e) must also be transferred to a digital accumulator. In a similar manner, the quantity y is then sampled and gated to the B-input of the multiplier, while the digital exponent, y_e, is transferred to the digital accumulator. The analog multiplier then forms the product xy at its output. Since the exponent of the product is equal to the sum of the exponents of x and y, the digital system must form

$$(xy)_e = x_e + y_e.$$

The sign of the product is also determined by the digital system.

All inputs to the multiplier are in the optimum range, but the output machine voltage may or may not be in the optimum range. Only two ranges are possible, however, for the following reason. Since

$$0.1 \leq x \leq 1.0 \quad \text{and} \quad 0.1 \leq y \leq 1.0,$$

then

$$xy_{\min} = 0.01 \quad \text{and} \quad xy_{\max} = 1.0.$$

Two ranges are therefore possible, and correction of the output machine voltage to the optimum range requires a decision element to determine whether the output voltage is initially in the range 0.01 to 0.1 machine unit or in the range 0.1 to 1.0 machine unit. To determine the range of the output voltage of the multiplier, the signal is fed to an analog comparator which triggers when the input voltage is greater than 0.1 machine unit. The output of the comparator, denoted F_m in Fig. 2, is then fed through a digital gate to the set input of a flip-flop. This is necessary since the digital system samples the value F_m at a particular time, and it is not desirable for this value to change thereafter. If the value did change after the digital system sampled F_m, the exponent would not be properly changed and an error would result.

Note that immediately before F_m is sampled, the flip-flop must be reset. If F_m is zero, the machine voltage must be multiplied by 10. Therefore, G_{63} is enabled by the output F_m of the flip-flop, and the voltage is fed through an operational amplifier of gain 10. If F_m is one, the machine voltage must be multiplied by 1. G_{64} is enabled by the output F_m, and the voltage is fed through the unity-gain channel of the operational amplifier. If the machine voltage is multiplied by ten, the digital exponent is also corrected.

The analog divider is provided for division of two analog voltages. The steps necessary for formation of the analog quotient and its associated scale factor are essentially identical to those necessary for multiplication. In this case, however, the digital system computes the difference in the exponents of x and y rather than the sum.

Because all input voltages to the divider are in the optimum range of 0.1 to 1.0 machine unit, the output voltage may lie between 0.1 and 10 machine units. This range extends outside the permissible machine voltage range of zero to one. Therefore, an internal fixed gain of 0.1 is inserted in the divider so that the output voltage range is 0.01 to 1.0.

The analog adder is provided for addition or subtraction of two analog voltages. It is also used to optimize the scale factor of an arbitrary machine variable. This later use is explained more fully below.

To form the sum $x+y$, the analog voltage x is first sampled and transferred to the A-input of the adder. The quantity y is then sampled and transferred to the B-input. Because the scale factors of two quantities to be added must be identical, it is necessary to first compare the digital exponents of the machine variables and then to adjust the scale factor of the smaller quantity to coincide with that of the larger so that

$$x_e = y_e.$$

The digital system computes the difference in exponents of the two inputs and uses this difference to determine the necessary attenuation of the machine voltage having the smaller scale factor.

The above scale factor adjustment may be understood with the aid of Fig. 2. The A- and B-inputs normally pass through the unity-gain channel of their associated scale factor devices. When an exponent adjustment is made, the digital system provides signals which gate the proper input through an attenuator.

Since only positive voltages are used in the machine, each quantity has an associated digital sign bit. This sign bit determines whether either or both adder inputs are passed through an inverting amplifier. Thus, an addition may be programmed, but if the sign bits of the two quantities to be added are different, one signal is passed through an inverting amplifier to produce a negative voltage input to the adder.

The output of the adder is passed through an absolute value device. This device is necessary because of the possibility that the adder output may produce a negative voltage. The digital system assumes that the sign of the sum is positive. If the output voltage of the adder is negative, this fact is sensed by a comparator which is referenced at zero volts, and the signal from the comparator is used to change the digital sign bit of the output.

The adder is also used to optimize the scale factor of an arbitrary variable. To do this, the variable is transferred to the A-input of the adder, and a zero-code (digital number representing zero) transferred to the B-input. The scale factoring devices at the output of the adder are then used to change the machine voltage to the optimum range.

Since at least one of the inputs to the adder is in the optimum range (unless both are identically zero), full precision is maintained for the addition of two machine voltages. Full precision is also obtainable in the subtraction of two machine voltages provided the difference is at least 0.001 machine unit. Because it is impractical to encode very small voltages, any adder output less than 0.001 machine unit is denoted by the zero code. This case arises as the result of subtraction of two quantities of very nearly equal value.

The block diagram of the components needed for optimizing the scale factor of the adder output is shown in Fig. 2. The signal from the output of the adder is fed to a range encoder. This is a voltage comparison device with digital outputs A, B, C, D. These outputs are equal to either one or zero as shown in Fig. 2, and are used to correct the digital exponent and the analog machine voltage. For example, if the range encoder outputs are A, B, C equal to one and D equal to 0, the machine voltage is in the range 1.0 to 2.0 machine units. The machine voltage must be multiplied by 0.1 and the digital exponent corrected. The encoder enables G_{71} and the signal is passed through the 0.1 gain channel of an operational amplifier.

If the zero-code appears at an input to the adder, G_{69} or G_{70} is enabled so that the analog voltage input is also zero. If both inputs to the adder are the zero-code, the digital system automatically generates the zero-code as an output. When the output voltage of the adder is less than 0.001 machine unit, the zero-code is generated at the output.

To perform subtraction in the machine, the same steps are used as for addition, with one exception. If it is desired to form the difference $x-y$, the quantity x is always transferred to the A-input of the adder, and the quantity y is always transferred to the B-input. Then depending upon the digital sign bits of x and y, the signal is fed either directly into the adder or first through an inverter. The digital logic makes the above decisions.

The Digital System

The floating-point system automatically scales the analog machine voltage to an optimum range during computation. The digital system stores and processes the scale factor (exponent) and sign of each variable in the machine. After each mathematical operation is performed in the computer or after a voltage in the analog arithmetic element is optimized, the digital system computes the new scale factor and sign.

The analog system and the digital system perform simultaneously during the course of a given operation. When an analog quantity is transferred to a computing component in the analog arithmetic element, its corresponding exponent and sign are transferred to an accumulator. While an output is being formed in one of the analog computing elements, the accumulator computes either the sum or the difference of the two exponents for multiplication and division, respectively. Before addition or subtraction of two voltages is performed in the analog arithmetic element, the accumulator compares the exponents of the two operands and automatically adjusts the smaller absolute quantity to have the same scale factor as that of the larger.

At the end of an arithmetic operation in the analog system, information concerning the level of the analog voltage is transferred directly to the accumulator. This information is used by the digital system to rescale the analog voltage into the optimum range and to either increase or decrease the digital exponent by the proper amount. Therefore, it can be seen that although there is no direct flow of data from the analog system to the digital system, some information is actually exchanged.

As in the analog system, operations are controlled in the digital system by the digital control element. The particular sequence of operations performed is determined by the program stored on the drum. From one to eight operations are performed in the digital system during the period of one operation in the analog system, and the digital control element therefore supplies high-frequency (800 kc) control pulses to various parts of the digital system. This will be discussed more fully in the next section.

The type of logical circuitry used in the digital system is the dc system which employs the pentode gate tube. This type of logic, used extensively in the Whirlwind I

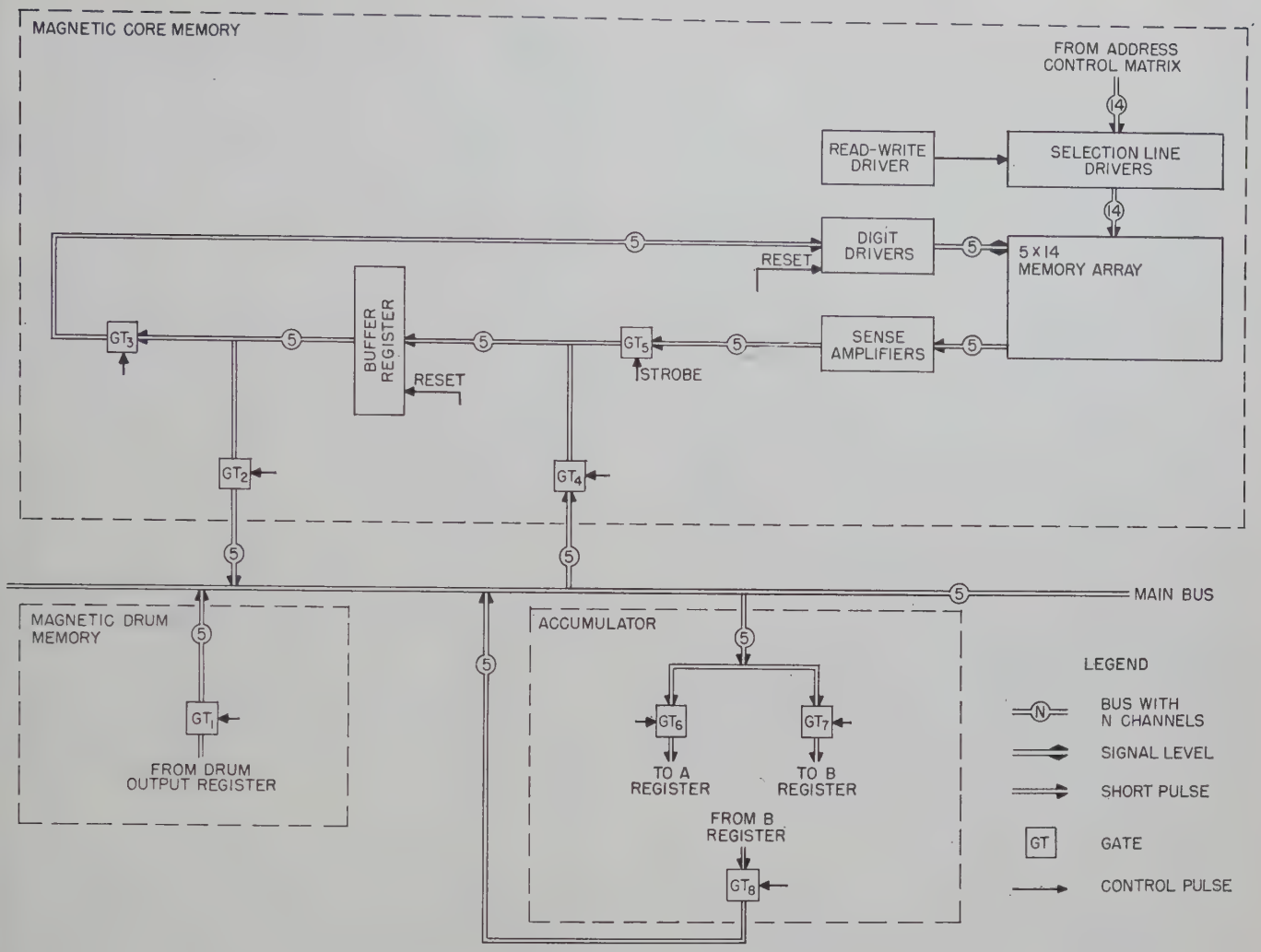


Fig. 3—Digital system.

computer[6] at the Massachusetts Institute of Technology is specifically designed for high-speed operation and has proved to be highly successful for real-time control. Systems of this type have been designed for pulse repetition rates up to two mc. The circuit logic is developed around the gate tube, the diode OR circuit, and the flip-flop. However, inverters, delays, diode AND circuits, and other such components are included in the system when necessary. This system of logic is also suitable for the use of transistors instead of vacuum tubes.[7]

The digital system, shown in Fig. 3, consists of the accumulator, the magnetic core storage, and the magnetic drum storage. Because digital words are transferred in parallel form, the main bus is composed of five parallel lines (four bits for exponent and one for sign). The information is transferred over the single bus set in binary form and in a synchronous manner. The transfers are executed by enabling digital gates connected to the main bus, as determined by the digital control element. As shown in Fig. 3, information is transferred to the main bus from one of three possible sources. Information can be transferred from the main bus, however, to one of only two possible destinations: the magnetic core storage or the digital accumulator. Only two destinations are possible because no information is written onto the magnetic drum during actual computation. In the analog system, information is always transferred from a storage unit to an arithmetic element, never from one storage unit to another. Consequently, no transfer of digital information from one magnetic core memory location to another is necessary. Such an order was intentionally avoided because it would require two digital addresses for one instruction.

A small-capacity memory is needed to store the exponent and sign associated with each machine variable since no information is written onto the magnetic drum during computation. The capacity will be approximately fourteen five-bit words, or a total of seventy bits. Since the outputs of the digital control element specify the address of the word, no register selection is

[6] M. F. Mann, R. R. Rathbone, and J. B. Bennett, "Whirlwind I, operation-logic," Digital Computer Lab., M.I.T., Rept. R-221; May 1, 1954.

[7] W. A. Clark, *et al.*, "The Lincoln TX-2 computer," Lincoln Lab., M.I.T., Memo. 6M-4963; April, 1957.

needed in the memory itself. The access time must be in the order of a few microseconds and random access is required.

Since the read process is essentially one of register selection, a one-dimensional read system can be used. A common read-out wire is used in each digit column. For writing, an additional dimension is needed for place-selection. Since the read-out process destroys the information held in the cores (*i.e.*, all cores in the register are set to the zero state), a regeneration cycle is needed to read out the information in a given register and write the same information back into the same register.

The digital accumulator performs all operations on the digital exponent and sign associated with a machine variable. When the proper gates are enabled, information is transferred from the main bus to either the A register or the B register in the accumulator. In all following examples, it will be assumed that the proper information is already in these two registers.

Five binary bits are required to represent the exponent and sign of a machine variable. Table I lists examples of scale factors and the digital numbers associated with them. The sign bit of the machine variable is not shown. The first digit column on the left represents the sign of the exponent and the next three digit columns represent the exponent. The one's complement is used to represent negative digital numbers.

TABLE I

Scale factors of machine variables	Digital representation
10^3	0 011
10^2	0 010
10^1	0 001
10^0	1 111
10^{-1}	1 110
10^{-2}	1 101
10^{-3}	1 100

The operations performed in the digital system during the multiplication of two machine variables will be explained in detail to illustrate a typical sequence of operations in the accumulator. When two machine variables are multiplied, the digital exponents of the variables must be added. In the accumulator the contents of register A are added to the contents of register B and the result stored in register B.

Two other functions must be performed in conjunction with this addition. First, the system must check to see if the zero-code appears in either register. A zero-code at either input means that the zero-code must be generated as an output. Second, the state of the analog comparator at the output of the multiplier in the analog arithmetic element must be sensed. The output of the comparator, denoted F_m in Fig. 2, determines the necessary correction in the digital exponent during the optimization of scale factors. The change in the digital exponent is

$$(\Delta e)_m = 0\ 00\overline{F}_m$$

and is added to the B register of the accumulator (which stores the results) to produce the corrected scale factor.

The functions that are performed by the accumulator during division of two machine variables are similar to those performed during multiplication. The exponent of the denominator, which is always stored in the B register, must be subtracted from the exponent of the numerator.

For the addition of two machine variables, the exponents of the addend and the augend must be identical. The difference in exponents is calculated and the exponent of the smaller is changed to be equal to that of the larger. The analog voltage associated with the smaller scale factor is then corrected. The sign bits of the machine variables are used in the analog arithmetic element to determine whether the machine voltages are passed through the inverters at the input of the adder.

Digital Control Element

The digital control element, shown in Fig. 4, controls simultaneous operations on the analog and the digital data. The main elements of this part of the system are the time pulse distributor, the operations control matrix, and the address control matrix. Inputs to these units (with the possible exception of the time pulse distributor) originate from the drum and are fed in from a drum output register.

From one to eight digital operations must be performed during the basic 10 μsec clock interval, and the time pulse distributor is provided to supply high-frequency control pulses to the digital system. This unit, which consists of a three-bit counter supplied by a synchronized 800-kc clock and a diode selection matrix, gates each of the eight high-frequency pulses onto one of its eight output lines. These eight lines, together with one of the output lines from the operations control matrix, send the specified sequence of from one to eight high-frequency control pulses to the digital system.

The operations control matrix and the address control matrix control the transfer of both analog and digital information in the machine. Two types of analog data transfers are possible: from analog storage to the input of a computing element (Type I), or from the output of a computing element to analog storage (Type II). Similar types of transfers occur in the digital system: from core or drum storage to the digital accumulator (Type I), or from the accumulator to core (but not drum) storage (Type II). Hence, two control matrices are used to control data transfers. For the Type I transfers the address control matrix specifies the origin of both the analog data (*i.e.*, the analog gate to be enabled) and the digital exponent (*i.e.*, the memory address of the stored number). The operations control matrix specifies the destination of the analog data (*i.e.*, the operation to be performed) and selects the appropriate group of digital control pulses for the specified operation. For Type II transfers, on the other hand, the address control matrix specifies both the destination of

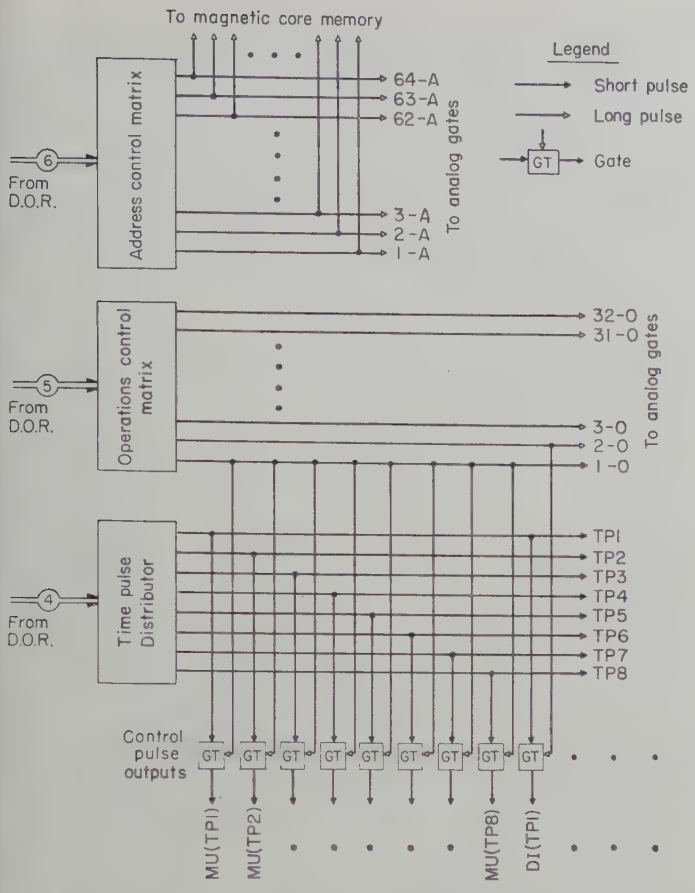


Fig. 4—Digital control element.

the analog data and the address in core memory in which the corresponding digital exponent is to be stored. The operations control matrix specifies the origin of the analog data (*i.e.*, the output of a computing element) and selects the group of digital control pulses necessary to transfer the digital word from the accumulator to the core memory location specified by the address control matrix.

The use of two control matrices allows the origin and the destination of data to be independently specified in a given instruction. If one control matrix were used, a separate instruction would have to be provided for every necessary combination of the two, so that the two-matrix system results in a much simpler order code. Furthermore, if use of a rectangular diode matrix is assumed, 544 diodes are necessary to implement both the six-bit address control matrix and the five-bit operations control matrix. At least ten bits would be required for a single control matrix, and implementation of this matrix would require 10,240 diodes.

Magnetic Drum Storage

The magnetic drum, which stores the program and input data, supplies timing pulses for the entire system and is the central element of the computer. Because the maximum access time is equal to the scanning period to one digital cell rather than to one period of drum rotation, it may be made comparable to that of a random access core memory; the response of the analog computing elements, however, limits the scanning period of 10 μsec.

Information is stored on the drum in parallel form in groups of tracks referred to as channels. Data words are recorded in two channels, one for the fractional part (eight tracks) and the other for the exponent and sign (five tracks). Five tracks constitute the instruction channel which feeds the operations control matrix, and six tracks constitute the instruction channel which feeds the address control matrix. Single tracks are provided for clock and origin pulses, and one track is used to supply positional information for function generation. The 800-kc clock for the digital system may be derived from high cell-density tracks on the drum or from a synchronized external source. Different channel capacities may be used if necessary but will remain fixed for a given computer.

The number of cells, D, around the periphery of a drum of diameter d is

$$D = cd$$

where c is the number of cells per unit length of track. If the computation cycle of an N-word program is to be completed in one drum revolution, the following inequality must be satisfied:

$$D \geq N.$$

The program length for one drum revolution is therefore limited by the drum diameter and the cell density. If longer programs are encountered, another set of instruction and data channels, together with appropriate switching arrangements, must be provided.

The rotation period of the drum determines the number of computation cycles per second, assuming that the program is completed in one drum revolution. If the minimum number of solutions required per second is S, then

$$S \leq \frac{1}{T} = S_c$$

where

T = rotation period of drum in seconds per revolution
S_c = number of complete computer solutions per second.

If, for example, a set of equations are to be solved in one drum revolution at the rate of fifty solutions per second, T must be made 1/50 second corresponding to a drum speed of 3000 rpm.

The following relationship must also be satisfied:

$$D = RT$$

where R is the scanning rate in cells per second. The choice of the three quantities in the above equation is not completely arbitrary. For real-time simulation problems, there is a definite minimum number of solutions per second required for accurate computation. For other simulation problems, however, the value of S may not be precisely defined, so that after the values of D

and R are chosen, the value of T can be calculated from the above equation. It should be emphasized that real-time simulation generally requires high-speed operation, so that if the values of D, R, and T cannot be chosen to make the above equation hold true, the computer cannot be used for the problem.

If it is desired to obtain solutions in more than one drum revolution, interlace methods could be used. Since this would require additional switching circuits, however, the interlace method would only be used to obtain more convenient values of D, R, and T. The equations presented above would have to be adjusted depending upon the type of interlacing used.

Programming

The basic programming techniques will be explained by an example of the multiplication of two machine variables, x and y, and the necessary instructions will be listed below. The instruction word consists of an operation and an address.

1) TMA x.
2) TMB drum.
3) MUL.

The first instruction (TMA) transfers the machine variable x from an analog storage device to the A-input of the multiplier and transfers the exponent x_e from core memory to the A register of the accumulator. The second instruction (TMB) transfers the variable y (assumed to be stored on the drum) to the B-input of the multiplier and transfers y_e from the drum to the B register of the accumulator. During the next order (MUL), the exponents are added and stored in the B register of the accumulator and the analog product is formed. Next, the state of the comparator at the output of the multiplier is sensed and the quantity $(\Delta e)_m$, which determines the necessary correction in the digital exponent during the optimization of scale factors, is subtracted from the B register. Thus, at the end of this order, the optimized analog product xy is at the output of the multiplier and its associated exponent $(xy)_e$ is stored in the B register of the accumulator.

It should be emphasized that the programmer is only concerned with the above orders and not with the operations or logic of the machine. He must only realize that the first two instructions are the memory access orders used for multiplication and that the MUL order is provided for multiplication and for scaling of the result. Programming of the equations for the real-time simulation of a physical system would be carried out in a similar manner.

System Evaluation and Conclusions

The analog-digital computer will be evaluated in terms of the class of problems that can be solved in real time, the complexity of the system, and the effort required to prepare a problem for computer solution.

At present, it is possible to operate the computer at a basic clock rate of 100 kc. At this frequency, addition and subtraction, including memory access time for both operands and scaling of the result, are performed in 50 μsec and multiplication and division in 30 μsec. The speed of the analog computing elements limits the solution rate of the computer at the present time, since digital systems have been operated at least an order of magnitude faster than in the present system. This indicates a need for further development of accurate high-speed analog computing components. To overcome the present speed limitations, it is conceivable that two or more arithmetic elements may be operated in parallel.

In a sense, the analog-digital computer sacrifices the precision of digital computation for the simplicity of analog components. The floating-point analog feature should, however, improve the accuracy. It should also be remembered that the computer was designed for simulation problems that require only moderately accurate solutions.

The use of time-shared analog computing elements in the system results in a major saving of equipment. The simplicity of the analog arithmetic element is one major advantage of the computer. Implementation of the arithmetic element of this computer with all-digital components would require more equipment, resulting in a greater total expense.

The use of an analog arithmetic element also provides easy access to the computer. Computers used for simulation, in most cases, must accept signals that are in the analog domain. Since the arithmetic element of the computer samples analog inputs and computes analog outputs which in both cases have fixed scale factors, no conversion equipment is necessary.

The floating-point analog feature requires additional equipment and so increases the system complexity. The digital system that is needed, however, is relatively simple. The word length is only five bits, and the digital elements must perform essentially only addition, subtraction, and comparison. The other functions of the digital system are storage and transfer of information. The advantages of the floating-point feature are increased dynamic range of variables, more effective use of the analog components, and ease of programming. Whether or not the advantages of the floating-point feature offset the additional equipment required will depend upon the particular application.

Scale factoring is a major problem in programming most present day computers, both analog and digital. The problem presented to the programmer is to keep all data, intermediate results and final answers within the computer's dynamic range. This can be a very involved process in the simulation of a complex physical system. With the analog-digital computer, however, the floating-point system automatically scales the machine variables to keep them within the computer's range. The programmer need consider only the maximum values of the variables. Hence, the programming of the

analog-digital computer is simpler and faster than that of purely analog simulators and all-digital simulators without special floating-point arithmetic.

One important advantage of the analog-digital computer is that the programming techniques can be easily learned. No knowledge of programming a general-purpose digital computer is required. By use of sets of instructions for the basic mathematical operations and subroutines for function generation, the equations can be programmed in the straightforward manner of hand-calculations. It should be emphasized that no knowledge of the computer logic is necessary to prepare a program for the solution of a new problem. Thus, engineers who are directly interested in simulation will be able to prepare their own programs for computer solution.

Based upon the logical design of the computer as described in this paper and upon the performance of a simplified analog arithmetic element now in operation, the following conclusions have been reached:

1) The magnetic drum is an effective means of information storage for this computer.
2) The use of time-shared analog computing elements results in a major saving of equipment.
3) The use of a high-speed analog arithmetic element provides easy access to input and output signals in the analog domain.
4) The floating-point analog feature increases the dynamic range of the computer, reduces scaling problems in programming, and allows more effective use of the analog elements.
5) The analog arithmetic element results in simpler construction but reduced accuracy as compared with a digital arithmetic element.
6) Digital control of analog computing elements provides an easy means of altering or completely changing a problem.
7) The programming techniques for this computer are easily learned and applied.
8) Computing in the combined analog-digital domain results in moderately accurate solutions. Increased accuracy can only be attained by the development of more accurate high-speed analog computing elements.

Appendix I

Analog integrators are currently being used to perform integration in the existing system. These units cannot be time-shared, and a separate unit therefore must be provided for each computed derivative.

Both positive and negative voltage inputs must be handled by the electronic integrators. Therefore, the sign bit of the computed derivative determines whether or not the input signal is passed through an inverter. For example, if $\dot{p}$ has a positive sign, it is stored through G_{16} and G_{17} of Fig. 2. The network consisting of G_{15}, G_{16}, and G_{76} and the inverter is time-shared among the inputs to all integrators.

The scale factor of an integrator output is fixed and, in general, is not optimum. Also, the polarity of the output voltage may be either positive or negative. Consequently, in this case negative voltages or machine variables with non-optimum scale factors must be transferred. Thus, when an integrator output is used in computations in the machine, corrections must first be made to optimize the scale factor and to change the voltage polarity if it is negative.

Analog integration has the advantage of furnishing continuous, smoothed outputs. The use of these analog devices suffers the disadvantage, however, of requiring a separate computing element for each derivative. The fact that the inputs and outputs of the analog integrators do not have optimum scale factors and may be either positive or negative is a further disadvantage. It should also be realized that integrators used in the manner described provide a rectangular approximation to the true integral which, in many cases, necessitates a relatively high solution rate to assure stability and accuracy of the solutions obtained.[8]

Because of the disadvantages of analog integration, it may be desirable to use numerical integration techniques. The analog integrators would then no longer be necessary, but analog storage devices would have to be provided to store the required number of past samples of the variable and its derivative. For example, if the rectangular rule were used, the following difference equation would be programmed:

$$Y_i(n) = Y_i(n-1) + T\dot{Y}_i(n)$$

where T is the sampling interval. One analog storage unit would be required to store $Y_i(n)$ between samples. Higher order integration rules may, of course, be used if additional storage units are provided. Two significant advantages arise through use of this method. Namely, quantities are handled with optimum scale factors, and at least part of the equipment involved can be utilized for other computations.

Appendix II

Arbitrary functions of one or more variables may be generated by the normal arithmetic operations of the computer if polynomial approximations are used. In many cases, however, it is desirable to approximate functions by straight-line segments and to linearly interpolate between stored discrete values of the function. If this method is to be used, discrete values of the function and the independent variable must be stored on the magnetic drum.

For example, if the function to be generated is $f(x)$, breakpoints of the independent variable (x_1, x_2, $\cdots$, x_i, x_{i+1}, $\cdots$, x_n) are stored on the drum. Following these numbers, the corresponding values of the function

[8] O. W. Kennedy, "Experimental analog-digital flight simulator," Electrical Engineer's thesis, Elect. Engrg. Dept., M.I.T.; January, 1959.

are stored $(f_1, f_2, \cdots, f_i, f_{i+1}, \cdots, f_n)$. Then, if $x_i \leq x \leq x_{i+1}$, the computer must select from the drum and temporarily store the following quantities: x_i, x_{i+1}, f_i and f_{i+1}.

The selection from the drum of the proper breakpoints and discrete values of the function is a time-selection problem which requires special logical equipment. This special logic has been designed but, because of its somewhat complex character, is not included in this paper. The method used in the data selection is as follows. The current machine variable x is compared with the stored breakpoints, and the special equipment gates x_i and x_{i+1} from the drum to S_1 and S_2, respectively, in Fig. 2. A counter, together with a single position track on the drum, is then used to determine the correct instant to read and store the numbers f_i and f_{i+1} in S_3 and S_4 respectively. It should be noted that as the fractional parts of the numbers are decoded and stored in analog storage units, their corresponding exponent and sign are automatically transferred from the drum to the core memory. After the four quantities are selected from the drum, the analog arithmetic element is used in the normal manner to solve the interpolation equation

$$f(x) = \frac{x - x_i}{x_{i+1} - x_i} [f_{i+1} - f_i] + f_i.$$

The special selection logic has been designed so that the above techniques are easily extended to the generation of functions of two variables. Generation of the function $f(x, y)$ requires the selection and storage in S_1 through S_8 of four breakpoints and four discrete values of the function.

Appendix III

The accuracies obtainable in the analog arithmetic element are largely determined by the characteristics of the gating and storage elements. Since the basic clock frequency is 100 kc, analog voltages must be sampled and stored in somewhat less than 10 μsec. The analog voltage level must then be accurately stored for a maximum period of one computation cycle. A typical solution rate is 50 computation cycles per second, so that the maximum holding time is 20,000 μsec. Assuming that a capacitor is used to store the analog voltage, the holding-to-charging time ratio is 2000:1. The desirable characteristics of the combined sampler and storage unit are:

1) The charging of the capacitor must be completed during the sampling interval.
2) The capacitor voltage must not decay appreciably during the holding time between sampling instants.

To analyze the errors of storage due to incomplete charging of the capacitor and decay during storage, refer to Fig. 5. The voltage is an exponentially rising waveform until the end of the sampling operation at

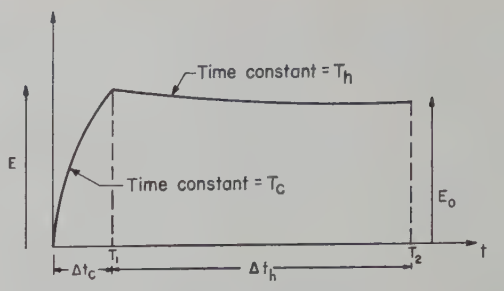


Fig. 5—Typical output of sampling and storage element.

time T_1. The voltage then decays exponentially during the holding interval Δt_h. The output voltage at the end of the holding interval is

$$E_0 = E[1 - e^{-\Delta t_c/T_c}][e^{-\Delta t_h/T_h}]$$

where

Δt_c = sampling interval
T_c = charge time-constant
E = step input voltage = asympotic value of the response during Δt_c
Δt_h = holding interval
T_h = storage time-constant.

The per unit change of output, defined as dE_0/E_0, is therefore

$$\frac{dE_0}{E_0} = \frac{dE}{E} + \frac{e^{-\Delta t_c/T_c} d(\Delta t_c/T_c)}{1 - e^{-\Delta t_c/T_c}} - d(\Delta t_h/T_h).$$

The condition for linearity is that

$$\frac{dE_0}{E_0} = \frac{dE}{E}.$$

The nonlinearity error, expressed in volts, is then

$$\epsilon = E \frac{e^{-\Delta t_c/T_c}}{1 - e^{-\Delta t_c/T_c}} d(\Delta t_c/T_c) - E d(\Delta t_h/T_h).$$

The nonlinearity error depends upon the magnitude of $\Delta t_c/T_c$, the change in $\Delta t_h/T_h$, and the change in the input voltage between sampling times. The error, however, does not depend upon the magnitude of Δt_h if T_h is a constant. The quantities Δt_c, T_c, and T_h may, for all practical purposes, be considered as constants, so that the nonlinearity error reduces to

$$\epsilon = -\frac{E}{T_h} d(\Delta t_h).$$

For all storage devices, except those connected to the integrator inputs, Δt_h is not equal to a constant. If the holding time were a constant, then the output of the storage device would be a linear function of the input voltage, and the error due to capacitor discharge could be corrected. The difficulty in storing the analog voltages is, therefore, not the fact that the holding time is large, but that the holding times vary among the analog storage units.

The following six papers were among those presented at the National Simulation Conference, Dallas, Texas, October 23–25, 1958. Three papers in the March, 1959 issue (those by D. W. Ladd and E. W. Wolf; Eugene Rawdin; and A. A. B. Pritsker, R. C. Van Buskirk, and J. K. Wetherbee) were also from that Conference, and we regret that they were not so identified. No National Simulation Conference is planned for this year.—*The Editor.*

Distributed Parameter Vibration with Structural Damping and Noise Excitation*

R. V. POWELL†

Summary—**A method is described for the electronic analog computer that will permit the determination of the vibration amplitude responses of a distributed system with structural damping to a random-noise excitation such as might be experienced by a missile structure accelerated by a jet propulsion system. There is general agreement among the investigators in the literature that structural damping is both frequency independent and amplitude dependent. Simulation of the structure by a method of normal modes permits the introduction of a discrete equivalent viscous-damping coefficient for each mode frequency, thereby effecting the frequency-independent characteristic of structural damping.**

Introduction

THIS paper describes a suitable analog representation of a distributed mechanical system, considering structural damping. The analog representation permits the determination of the displacement response of the system to an arbitrary forcing function. More specifically, however, the paper is concerned with the determination of the mean-square response of a structurally damped system to a random excitation on the electronic analog computer.

The electronic analog representation is made utilizing a method of normal modes. The method of normal modes used requires a system of computer amplifiers for each mode frequency considered. An alternative analog might be synthesized by finite-difference methods, resulting in a system of computer amplifiers for each system coordinate considered. However, it is apparent from the following considerations that the first alternative is the better one for this application. The normal-mode approach permits the introduction of a discrete damping coefficient for each mode frequency, while the finite-difference approach permits the introduction of a discrete damping coefficient for each coordinate. Considering, for example, the general form of a missile structure to be a long tube of uniform construction supporting internally unconnected masses suggests a system with damping reasonably independent of the system coordinate. Empirical data also indicate an amplitude dependence and a frequency independence of the structural damping coefficient. The normal-mode approach permits the introduction of an equivalent viscous damping which can be weighted to yield a frequency-independent damping coefficient. It should be noted, however, that this approach requires the damping coefficient to be equal for all coordinates of the system, but, as suggested previously, this is not a serious disadvantage for a missile-type structure.

Details of the analog representation of structural damping, the normal-mode method, and the computer program are presented.

Representation of Structural Damping in the Analog System

Structural damping is of interest in the vibration problem because it is the mechanism by which energy is dissipated in the vibrating system. As such, it is a determining factor in both the transient and the steady-state behavior of the system. As referred to in this paper, structural damping is a total damping capacity of the structure and, in general, is the result of several sources of damping, the more important of which are a material damping or solid friction and a joint or slippage damping such as is encountered in riveted joints.

Structural damping (sometimes called material damping) has been treated in the literature.[1–3] At present, however, there appears to be no suitable theory to describe the damping behavior of materials. Seitz[4] has

* Manuscript received by the PGEC, October 22, 1958; revised manuscript received, February 6, 1959. This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. DA-04-495-Ord 18, sponsored by the U. S. Army, Ordnance Corps. This paper was presented at the National Simulation Conference, Dallas, Texas, October 23–25, 1958.

† Jet Propulsion Lab., California Inst. Tech., Pasadena, Calif.

[1] J. M. Robertson and A. J. Yorgiadis, "Internal friction in engineering materials," *J. Appl. Mech.*, vol. 13, pp. A173–A182; September, 1946.

[2] A. L. Kimball, "Vibration problems—friction and damping in vibrations," *J. Appl. Mech.*, vol. 63, pp. A37–A41, A135–A140; 1941.

[3] B. J. Lazan, "Effect of damping constants and stress distribution on the resonance response of members," *J. Appl. Mech.*, vol. 20, pp. 201–209; June, 1953.

[4] F. Seitz, "Physics of Metals," McGraw-Hill Book Co., Inc., New York, N. Y., p. 67; 1943.

listed as possible sources of energy dissipation the following: ferromagnetic effect, plastic flow effect, and intergranular thermal currents. Consideration of these effects does not lead to a theory capable of predicting the behavior of materials in vibrating systems under normal engineering stress levels.

Among the investigators in the literature there is general agreement that the damping capacity of a material is independent of the frequency of vibration and is a function of the stress amplitude. The results of many of the investigations suggest the following empirical relation,

$$D = JS^n \tag{1}$$

where D is the damping capacity of the system defined as the energy lost per cycle of vibration, S is the stress amplitude, and J and n are constants associated with the material under study. For most materials the exponent n is between 2 and 3, but has been reported in excess of 20 for certain materials.

A specific damping capacity is usually defined as the ratio of the energy loss per cycle to the energy in the system,

$$\psi = \frac{\Delta w}{w} . \tag{2}$$

The specific damping capacity corresponds to the logarithmic decrement familiar to viscous damping, which is the ratio of amplitudes at the beginning and end of a complete cycle. Since energy in the system varies as the square of the amplitude, it follows that a logarithmic decrement is useful only for that case in which n is equal to 2. Material damping is, in general, described by (1) with n greater than 2, resulting in a decrement per cycle which is not logarithmic.

Before discussing a representation of damping it might be well to look at the problem of determining the damping capacity of a system from experiment. The physical results of material damping are evident as a temperature rise during the start of vibration, a thermal equilibrium in the steady-state vibration, a hysteresis loop in the stress-strain plot, a finite resonance amplitude, a wider resonance curve, and a decremental decay of amplitude per cycle for a freely vibrating system. In theory, damping can be determined by measuring any of the preceding, and the results can be expressed as a damping capacity. The preceding measurements are not all equally practicable at stress levels commonly encountered in engineering practice; however, a more fundamental objection to an arbitrary selection results from a consideration of the analog representation. If the analog model is sufficiently exact to represent each of the above physical phenomena with equal fidelity, the choice is independent of the analog. In general, most models will be inexact to a greater degree for some of the phenomena than for others, making the phenomenon of greatest interest a determining factor in matching the analog to the physical system.

Complex damping has enjoyed some popularity for describing structural damping coefficients. This method results in a frequency-independent damping coefficient, but it neglects the amplitude dependence of the coefficient. The form of complex damping as used by Theodorsen[5] and others in flutter studies may be expressed by writing the equation of motion as,

$$m\ddot{x} + k(1 - ig)x = F(t). \tag{3}$$

The component $igkx$ is seen to be in phase with the velocity and constitutes the damping force. As pointed out by Soroka,[6] the representation of (3) has some disadvantages, a particular disadvantage being that the resonant frequency increases with increased damping rather than decreasing as is verified experimentally. A better representation of complex damping is offered by Myklestad.[7] Here, the equation of motion is written,

$$m\ddot{x} + ke^{2bi}x = F(t). \tag{4}$$

The damping is taken care of by writing the complex spring constant ke^{2bi}. The effective spring force is now $kx \cos 2b$, and the damping force is $kx \sin 2b$. The coefficients g in (3) and b in (4) are simply related to the per unit critical damping of viscous damping. Complex damping can be modeled on the electronic analog computer by resorting to frequency-independent phase shifters. Such a method limits the frequency range to that of the phase shifters, and this may be a considerable restriction in practice.

A more exact model may also be considered to represent a damping such as expressed by (1). A reasonable approximation can be achieved by a diode simulation of the hysteresis loop; however, it is not the opinion of the author that the introduction of these complications contributes to a more exact solution.

In a problem of stress or displacement response, it is probably the shape of the resonance curve or the resonance amplitude which is of most interest. For small damping, the resonance curve at any single frequency of the structurally-damped system deviates very little from the resonance curve of a viscously-damped system for a suitable choice of equivalent viscous damping. The equivalent viscous damping may be chosen for either an amplitude match or a bandwidth match between the analog and the physical system. Since the normal-mode approach permits separate coefficients for discrete frequencies, a good approximation to the total system can be realized.

[5] R. Rosenbaum and R. H. Scanlan, "Aircraft Vibration and Flutter," The Macmillan Company, New York, N. Y., p. 67; 1951.
[6] W. W. Soroka, "Note on the relations between viscous and structural damping coefficients," *J. Aero. Sci.*, vol. 16, p. 409; 1949.
[7] N. O. Myklestad, "The concept of complex damping," *J. Appl. Mech.*, vol. 19, p. 284; September, 1952.

It should be noted that the normal-mode analog is not an ideal representation of (1), since there is no amplitude dependence of the damping. Nonlinearities may not be introduced into the normal-mode system; however, an iterative approach to a better damping coefficient choice might prove fruitful.

Electronic Analog Representation of a Structure by Normal Modes

By the method of normal coordinates the displacement response of a linear conservative distributed system can be determined from the normal-mode shapes and the mass distribution of the system.[8] The displacement can be expressed as the sum of the mode displacement responses

$$u(x, y, z, t) = \sum_n q_n(t)\phi_n(x, y, z)$$

where the q_n's are the solutions of the Lagrange equations

$$\ddot{q}_n + \omega_n^2 q_n = \frac{Q_n(t)}{M_n}. \tag{5}$$

$Q_n(t)$ is the generalized force and is given by

$$Q_n(t) = \int f(x, y, z, t)\phi_n(x, y, z)dv,$$

which also can be expressed

$$Q_n(t) = \sum_m \phi_{mn} f_m. \tag{6}$$

M_n is the generalized mass and is given by

$$M_n = \int \phi_n^2(x, y, z)dm$$

which, if coupled mass terms are neglected, can be expressed as

$$M_n = \sum_m \phi_{mn}^2 M_m \tag{7}$$

where ϕ_{mn} is the mode shape of the mth coordinate associated with the nth natural frequency, f_m is the arbitrary forcing function, acting on the mth coordinate of the system, and M_m is the mass distribution of the system.

The solutions of the Lagrange equations are given by the Duhamel integral as

$$q_n = \frac{1}{M_n\omega_n}\int_{-\infty}^{t} Q_n(\tau) \sin \omega_n(t - \tau)d\tau. \tag{8}$$

The displacement response of the system is then given by superposition of the mode responses.

$$U_m = \sum_n q_n(t)\phi_{mn}. \tag{9}$$

Strictly speaking, the nonconservative system can be reduced to normal coordinates only if the retarding force due to damping takes a particular form, *i.e.*, a retarding force acting on each part of the system proportional to both the mass and the velocity of the part. Rayleigh[9] points out, however, that for small damping the errors of the approximation afforded by reduction to a normal coordinate are of second order.

The normal-mode method may be extended to the nonconservative system by writing the mode equations in the form

$$\ddot{q}_n + 2\gamma_n\omega_n\dot{q}_n + \omega_n^2 q_n = \frac{Q_n(t)}{M_n} \tag{10}$$

where γ is an equivalent viscous damping.

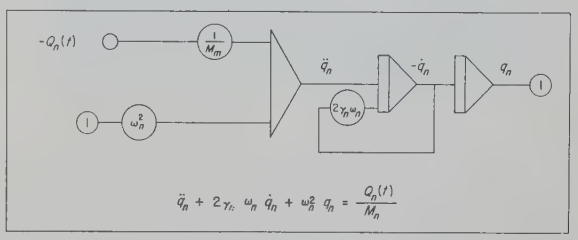


Fig. 1—Computer program for each normal mode.

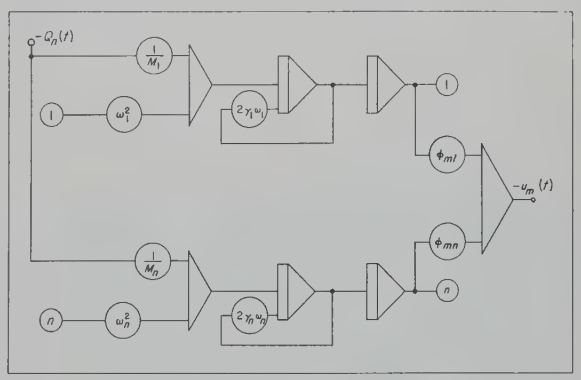


Fig. 2—Complete computer program.

Solutions of (8) are not readily available; however, solutions of the Lagrange equations (5), or the extended equation for an equivalent damping (10), are readily obtained on the electronic analog computer. A typical computer setup necessary for each of the normal modes considered is presented in Fig. 1. Summation of the mode displacements q_n to realize the system displacement u_n is also readily accomplished on the electronic analog computer. Fig. 2 illustrates a complete computer setup for determination of a system displace-

[8] T. Von Kármán and M. A. Biot, "Mathematical Methods in Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 162–196; 1940.

[9] J. W. S. Rayleigh, "The Theory of Sound," Dover Publications Inc., New York, N. Y., pp. 69–173; 1945.

ment response considering n normal modes. It should be noted that in practice the system response for various system coordinates is determined simply by resetting the coefficient potentiometers involving the mode shapes ϕ_{mn} and calculating a new $Q_n(t)$.

In summary, the displacement response of a linear distributed system with small damping may be obtained on the electronic analog computer, if the mass distribution, some reasonable number of mode shapes and mode frequencies, and an equivalent viscous damping for the whole system are known.

Instrumentation of the System Response

For the problem of random excitation, the mean-square response which is of interest[10,11] is defined by

$$\overline{x^2} = \frac{1}{T}\int_0^T x^2 dt. \tag{11}$$

Since the computer output is just x, the mean-square response may be obtained by squaring the response and integrating for a unit time. Integration of the squared output dictates some frequency-response restrictions on the problem setup. An upper limit is imposed on the integration time by the drift inherent in the integrating amplifiers and by the time required to obtain a single piece of data.

The minimum integration time is, in turn, dependent on the lowest frequencies in the response, so that a minimum frequency is established for the problem. The computing amplifiers are good to at least 1 kc, so the frequency restrictions are effectively a function only of the input and output systems. The input is a noise generator or equivalent and can easily be extended in frequency. In addition to the integrator, the output requires a wide-band squaring device.

[10] W. T. Thomson and M. V. Barton, "The response of mechanical systems to random excitation," *J. Appl. Mech.*, vol. 24, pp. 248–251; June, 1957.
[11] R. H. Battin and J. H. Laning, "Random Processes in Automatic Control," McGraw-Hill Book Co. Inc., New York, N. Y.; 1956.

Optimization by Random Search on the Analog Computer*

J. K. MUNSON† and A. I. RUBIN‡

Summary—**One method of searching a system for optimum operating conditions is to evaluate system performance for many randomly-chosen combinations of the independent parameters. This paper explains the use of standard electronic analog computer equipment to accomplish such a search of a mathematical model quickly and economically. Gaussian noise sources generate values of the independent parameters and sample-hold circuits hold those values which give the best value of the optimization criterion. An application of the method to a production allocation problem is mentioned.**

Introduction

RECENT years have seen increased emphasis on optimum design and operation of military weapons systems, physical and chemical processing systems, and business systems. Generally, such systems can be designed and operated under many different sets of conditions. The purpose of an optimization study is to determine the best combination of such conditions within restraints imposed by economy, physical size, and other limitations of a system or surroundings.

An investigation of a typical system includes the following steps:

1) Describe system operation by a set of equations obtained theoretically or through study of an existing pilot unit.
2) Specify restraints on the variable system parameters.
3) Express the characteristic to be optimized as a function of all significant system variables and define the "optimum" value.
4) Devise a method of choosing combinations of system parameters for evaluation which will allow selection of optimum conditions in an economical manner.

For example, consider a chemical reactor which will operate under many combinations of temperature, pressure, catalyst activity, and reactant concentrations. Optimum design and operation of this reactor is defined as that combination of conditions which will supply the available market demand at the lowest operating cost. In order to find the optimum, it is necessary to have available a model reactor or its mathematical equivalent. Restraints on the variables must be specified, for

* Manuscript received by the PGEC, November 17, 1958; revised manuscript received, April 3, 1959. This paper was presented at the National Simulation Conference, Dallas, Texas, October 23–25, 1958.
† E. I. Du Pont de Nemours and Company, Wilmington, Del.
‡ Electronic Associates, Inc., Princeton, N. J.

example, the maximum pressure allowed at any given temperature due to strength of materials or capabilities of auxiliary equipment such as pumps and blowers. An equation describing the operating cost as a function of the system variables must be written. A method of experimentation is necessary which will locate the area of optimum operation with an economical expenditure of time and effort.

The traditional role of the analog computer in optimization studies has been to act as a mathematical model of the system to be investigated. Design parameters and operating conditions can be easily changed and system operation determined quickly, economically, and safely. This paper describes an application of the analog computer which decreases the amount of statistical analysis and auxiliary calculation required. The high speed of the computer is utilized to evaluate many more combinations of operating variables than would otherwise be economically feasible.

Brooks[1] describes four basically different ways of conducting experiments in order to find a maximum, and compares random methods to the factorial, univariate and steepest ascent methods. One advantage quoted for the random method is that less trials may be required where there are a large number of variables involved. Brooks gives a quantitative analysis of the number of trials required for an idealized experimental situation.

Analog computer techniques have been applied to linear and nonlinear programming methods for optimization.[2,3] The random search technique described here will use only about one-half of the amount of analog equipment required in the linear programming method, but will, in general, take more time to find an optimum. Thus, the random search method allows solution of much bigger problems on existing analog installations.

Random Search Method

The basic principle of this random search for optimum is this. By specifying each of the parameters to be investigated in a random manner within the region of interest, there is a finite probability that the entire region will be searched and the optimum discovered in a reasonable time. The time required will depend on characteristics of the systems, the size and frequency of parameter variations, and the precision with which the optimum is desired.

Fig. 1 is a block diagram illustrating the method for a system described by a set of nonlinear algebraic equations. A set of random noise generators continuously generate values for each of the system parameters. Practical limits on the range of variation are imposed through a group of restraints which can be constants or

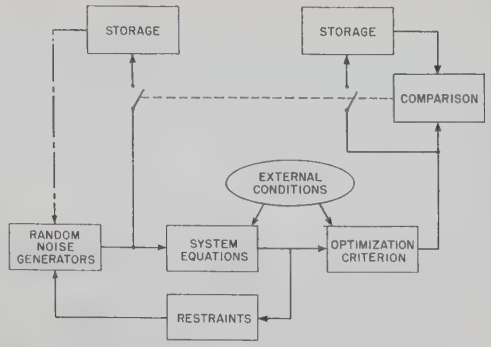


Fig. 1—Block diagram of random search method.

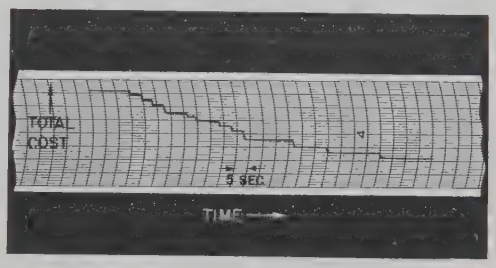


Fig. 2—Time plot of computer solution.

functions of any of the system variables. The outputs of the system are a function of both the particular values chosen for the operating parameters and the external conditions under which the system must operate. The optimization criterion describes how well the system operates with the set of parameters present at the input. This continuously varying value of the optimization criterion is compared with a previously-stored value. If, at any time, the instantaneous value is closer to a maximum (or minimum, depending on which is desired), this new value of the function and the set of system parameters producing it are switched to the storage circuits to replace the previous "best" values. After a period of time, determined by the characteristics of the system and the noise sources, the stored values will stop changing significantly and the optimum has been found.

The region of search may be narrowed through the use of the fact that the storage circuits contain the coordinates of the point closest to the optimum of all those previously evaluated. If further random search is concentrated around these points, we should reach the optimum more quickly. Since use of this procedure increases the possibility of locating a local optimum only, the search is repeated from several different starting points to see if the same point is reached by different paths.

For the application described below, Fig. 2 is a typical time plot of the optimization function.

This random search method could be followed in hand computation or by digital as well as analog means. However, since the analog is a parallel machine which runs continuously, many more input combinations can be evaluated in a given length of time by the analog. For

[1] S. H. Brooks, "A discussion of random methods for seeking maxima," *Operations Res.*, vol. 6, pp. 244–251; March–April, 1958.
[2] I. B. Pyne, "Linear programming on an electronic analogue computer," *Commun. and Electronics* (*Trans. AIEE*, pt. I, vol. 75), pp. 139–143; May, 1956.
[3] J. D. Strong, "Optimizing with a Computer," *Automatic Control*, vol. 4, pp. 16–19; June, 1956.

instance, consider a problem in which 1 per cent variations in the variables are significant. If the noise were filtered to a maximum frequency of 10 cycles per second, a maximum of 1000 values could be generated in one second by each noise source. Obviously as the number of variables to be generated in this way increases, the analog has a distinct advantage.

The explanation above applies to the exploration of systems described by a set of algebraic equations. With dynamic systems, the problem formulation requires a set of differential equations and the quantity to be optimized may be a function of the system's time history, such as the smallest integrated error under a given upset. In such a case the system parameters must obviously be constant during the period of time that the upset is to be applied. Favreau and Franks[4] explain this application of the random search method. The essential modification is a "hold" circuit which samples the randomly varying parameter signals and holds them constant during each solution of the differential equations.

Typical Circuitry

Random Noise Generators

Noise signals with a Gaussian distribution were available from an Electronic Associates Model 201A noise generator. The pass band of the generator is from dc to 30 cps. The source of noise is a gas thyratron.

An interesting possibility exists for decreasing the amount of special noise generation equipment required. If there is no special requirement on the noise characteristics, it should be possible to delay the output of one source through several channels with different delays to obtain a number of random signals.

Comparison—Storage Circuits

The circuit of Fig. 3 is used to compare the instantaneous value (E) of the optimization criterion with the best previously attained (E^*) and to switch this signal and the corresponding system inputs into the storage circuits if the new value is better.

If a maximum E is desired, E^* will usually be larger than E. The relay contacts will be connected as shown. When E becomes greater then E^*, the relay will change position and, with the component values shown, the output of the storage amplifier will attain the value of E present when the relay switched. This same relay amplifier will control similar storage circuits on each of the system inputs. When E^* has reached its new value, the relay returns to the position shown.

If the random search is to be concentrated around the point giving the best previous value, the random inputs will be added to the value being held on these input storage circuits.

[4] R. R. Favreau and R. G. E. Franks, "Statistical Optimization," presented at the Second International Conference for Analog Computation, Strasbourg, France, September, 1958.

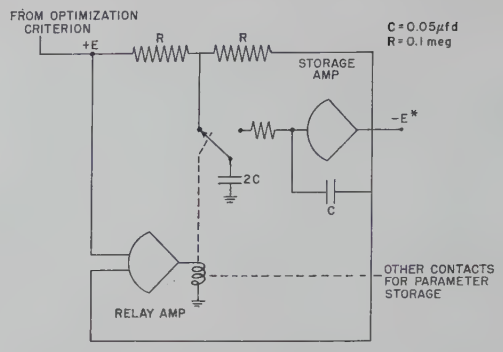


Fig. 3—Comparison—storage circuits.

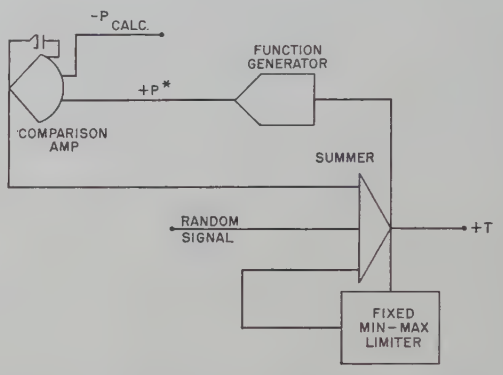


Fig. 4—Circuit for introducing system restraints.

Restraint Limiters

Restraints on the variations of the system operation are a necessary part of the statement of the optimization problem. In addition to the obvious limits like maximum physical capacity, there are limits on some quantities which are defined as a function of other system variables. As an example, consider the reactor mentioned previously. The limits on strength of the materials of construction define a maximum pressure for any given temperature. This restraint on the system must be introduced if both temperature and pressure are to be generated independently. Fig. 4 shows a means for performing this function. A random signal is available to define the temperature T. Conventional circuits establish minimum and maximum levels based on the region being searched. There exists a pressure limit (P^*) for any temperature. The pressure limit can be a nonlinear function of T. This P^* is compared with the pressure calculated for the system. If P_{calc} should attempt to exceed P^* an impractical solution exists and either temperature, pressure or both must be reduced. In this case, the circuit is connected to reduce T.

Application

The optimum operating conditions for a production system, a portion of which is shown in Fig. 5, was determined by the random search technique. In this system, two plants make several of the same products. Each of these products has a sales demand and is used as raw material for one or more other processes. Because

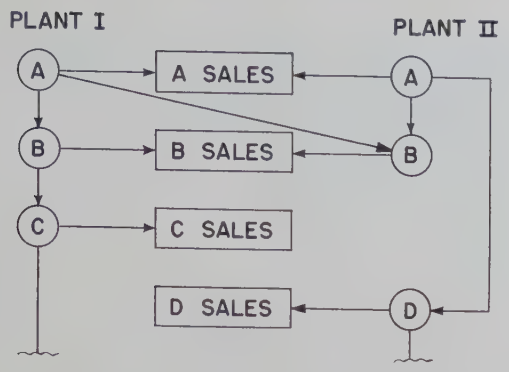


Fig. 5—Portion of production system studied.

Typical Equations

$P_1 = P_7 + P_9 + P_{11} + 0.632\ P_3 + 1.161\ S_7 + S_8$
$P_2 = S_1 + S_2 + S_3 - P_7 - P_9 - P_{11}$
$P_3 = \text{Random}$
$P_4 = S_{11} + 1.053\ S_6 - P_{13}$
.
.
.
$\text{Cost} = K_1P_1 + f(P_2) + f(P_3) + K_4P_4 + f(P_5) \cdots$

Typical Restraints

$0 \leq P_1 \leq S_1$
$0 \leq P_2 \leq 1.07\ C_9$
$0.637\ S_5 \leq P_3 \leq C_3$
.
.
.
.

$P = \text{Production rate}$
$S = \text{Sales demand}$
$C = \text{Maximum capacity}$

Fig. 6—Typical equations and restraints.

of geographical location, age of equipment, availability of labor and other similar factors, the costs of production at the two plants differ for each product. The problem was to determine, for any particular set of sales demands, the split of production between the two plants which would minimize the over-all cost.

The system was described by a set of 10 simultaneous algebraic equations in 15 unknowns. For example, the production of A at plant I is a function of not only the sales demand for A, but also is a function of the sales demands of B, C, and D and of the amounts of these products produced at plant II. In addition, a group of restraint equations define minimum and maximum limits for each variable.

Fig. 6 contains typical equations for the production system considered here. For instance, the first equation shows that the required production of product 1 is equal to the sum of several other production quantities and sales figures multiplied by appropriate stoichiometric constants. These constants show that one pound of product 1 is required for manufacture of products 7, 9, and 11, that 0.632 pound of 1 is required to make one pound of 3, etc. Typical restraint equations show that all production quantities must be positive and that some have a minimum as well as a maximum limit.

The criterion for optimization was an over-all cost

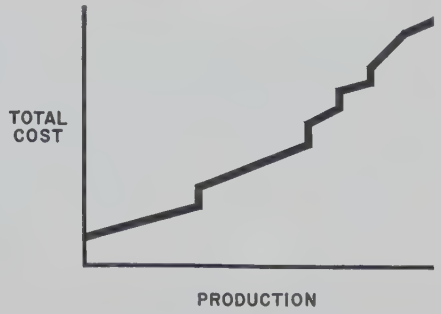


Fig. 7—Typical cost curve.

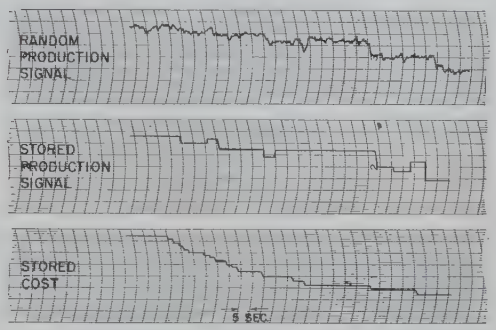


Fig. 8—Sample records.

equation representing the sum of the individual nonlinear cost functions. Fig. 7 is a typical cost curve showing the cost as a function of production rate. The sharp breaks in the curve show the necessity of hiring new personnel or utilizing more or different pieces of equipment. The slopes of the straight sections are not in general equal.

The computer circuitry used was conventional summing equipment for the linear algebraic relations of Fig. 6. Five noise generators were required along with the circuits of Figs. 3 and 4 in the complete system.

By using five random signal sources and directing the search around the best previous point, the optimum was located usually in less than a minute for each set of sales demands to be investigated. Typical strip chart recordings of several variables are shown in Fig. 8.

Conclusion

A method has been developed to utilize the high speed of the analog computer in conducting a random search for optimum conditions. This technique should be useful not only in optimization studies, but may find application in other similar problems. For instance, knowing the form of an equation or set of equations, it may be necessary to determine a group of constants from experimental data. Sets of algebraic equations are sometimes solved by assuming various values of the unknowns and searching for the set which will give the smallest set of residuals. It is possible that a computer-controlled process might be kept operating at optimum conditions by such random search techniques applied to the process itself or, more likely, a fast-time representation of it.

Linear System Approximation by Differential Analyzer Simulation of Orthonormal Approximation Functions*

ELMER G. GILBERT†

Summary—Various analytic procedures have been proposed for minimum integral-square-error approximation of prescribed linear systems; however, they often involve computational difficulties. In the procedure developed in this paper, a group of N linear approximating systems with orthonormal impulse responses $\phi_n(t)$ are realized by operational amplifier circuits. When $h(-t)$ forces the systems ($h(t)$ is the impulse response of the prescribed system) it is found that their outputs at $t=0$ are a_n, the coefficients in

$$h^*(t) = \sum_{n=1}^{N} a_n\phi_n(t),$$

the approximate impulse response. The following points relative to the approximation procedure are developed: constrained and weighted integral-square-error approximations, derivation and realization of orthonormal functions, physical realization of $h^*(t)$, evaluation of error $h(t)-h^*(t)$, and analysis of computer errors. Several approximation examples are given.

Introduction

THE approximation problem of linear system synthesis may be stated as the determination of a *realizable* system function which closely approximates a *prescribed* system function. Though various procedures have been developed for solving the approximation problem,[1] none yield simple calculations when the prescribed system is not specified in elementary analytic form. This is particularly true when the prescribed system function is the impulse response, $h(t)$, *i.e.*, the approximation is in the time domain. Some steps have been made to alleviate the computational troubles through the use of digital computers.[2,3] This paper presents a differential analyzer method for minimum integral-square-error approximation in the time domain. No difficult analytic computations are required and a simulation of the approximate system is obtained as a bonus.

The Orthonormal Function Approximation

The approximate impulse response is defined by a sum of predetermined approximating functions[4] $\phi_n(t)$ as

$$h^*(t) = \sum_{n=1}^{N} a_n\phi_n(t) \tag{1}$$

and is chosen (by varying the a_n) to make the integral square error,

$$E = \int_{-\infty}^{\infty}(h-h^*)^2dt = \int_{-\infty}^{\infty}h^2dt - 2\sum_{n=1}^{N}a_n\int_{-\infty}^{\infty}h\phi_n dt + \sum_{n=1}^{N}\sum_{m=1}^{N}a_na_m\int_{-\infty}^{\infty}\phi_n\phi_m dt, \tag{2}$$

a minimum. When the ϕ_n are orthonormal, that is, they satisfy the relation

$$\int_{-\infty}^{\infty}\phi_n\phi_m dt = 0, \quad n \neq m$$
$$= 1, \quad n = m \tag{3}$$

(2) becomes simply

$$E = \int_{-\infty}^{\infty}h^2dt - 2\sum_{n=1}^{N}a_n\int_{-\infty}^{\infty}h\phi_n dt + \sum_{n=1}^{N}a_n^2. \tag{4}$$

Setting $\partial E/\partial a_n = 0$ for all n to minimize E in (4) gives

$$a_n = \int_{-\infty}^{\infty}h\phi_n dt \tag{5}$$

and a minimum integral-square-error

$$E_{\min} = \int_{-\infty}^{\infty}h^2dt - \sum_{n=1}^{N}a_n^2. \tag{6}$$

The above steps require that $\int_{-\infty}^{\infty}h^2dt < \infty$; realizability of h^* demands that h^*, and consequently the ϕ_n, be zero for negative time. Note that $E_{\min}$ decreases as the number of terms, N, in h^* increases.

The transfer function of the approximate system is given by the Laplace transform of (1) as

$$H^*(s) = \sum_{n=1}^{N} a_n\Phi_n(s) \tag{7}$$

where the upper-case-letter functions of s are the Laplace transforms of the corresponding lower-case-letter functions of t. Since only lumped element realizations are considered in this paper,[5] $H^*(s)$ and the $\Phi_n(s)$ are rational in s.

* Manuscript received by the PGEC, September 22, 1958; revised manuscript received, February 13, 1959. This paper was presented at the National Simulation Conference, Dallas, Texas, October 23–25, 1958.

† Dept. of Aeronaut. Engrg., Univ. of Michigan, Ann Arbor, Mich.

[1] S. Winkler, "The approximation problem of network synthesis," IRE Trans. on Circuit Theory, vol. CT-4, pp. 5–21; September, 1954.

[2] G. A. Caryotakis, H. B. Demuth and A. D. Moore, "Iterative network synthesis," 1955 IRE Convention Record, pt. 2, pp. 9–16.

[3] M. R. Aaron, "The use of least squares in system design," IRE Trans. on Circuit Theory, vol. CT-3, pp. 224–231; December, 1956.

[4] Proper choice of the $\phi_n(t)$ improves approximation accuracy for fixed N. A procedure for improving on the initial selection of $\phi_n(t)$ will be indicated later.

[5] Distributed element realizations also have useful approximation properties. See forthcoming paper by E. G. Gilbert and J. Otterman: "The Synthesis of Linear Filters with Real or Imaginary Transfer Functions."

Orthonormal function approximations of the above type are well known[6] and have been employed for linear system synthesis by several authors.[7–9] Their main disadvantage lies in the difficulty of evaluating the a_n by (5). Analytic procedures are lengthy and are only straightforward when $h(t)$ has a known Laplace transform.[10] The differential analyzer evaluations of (5), which are discussed in the next section, require only that $h(-t)$ be available as a computer variable.

Since the integral-square-error criterion is essential to the orthonormal function approximation, it is important to discuss some of its properties.

Because of the square weighting large error magnitudes tend to be excluded by minimization of E. The result is an approximation error which generally oscillates about zero with relatively constant peak amplitude. Hence a good approximation is not confined to the region of a single point as in a Taylor series expansion. A disadvantage of the square weighting is that there is no time interval of appreciable length (at least where h or h^* have appreciable value) where the error is very small. This disadvantage can often be relieved by using a weighted integral-square-error approximation which is discussed in a later section.

Approximation in the time domain also brings about an approximation in the frequency domain; a fact which can be seen from an application of Parseval's theorem to (2).

$$E = \int_{-\infty}^{\infty} (h - h^*)^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} | H(j\omega) - H^*(j\omega) |^2 d\omega. \quad (8)$$

From (8) it is easily shown that

$$E_G = \frac{1}{2\pi} \int_{-\infty}^{\infty} (|H| - |H^*|)^2 d\omega \leq E, \quad (9)$$

$$E_R = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\mathrm{Re}\, H - \mathrm{Re}\, H^*)^2 d\omega \leq E, \quad (10)$$

$$E_I = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\mathrm{Im}\, H - \mathrm{Im}\, H^*)^2 d\omega \leq E. \quad (11)$$

Thus a good approximation of $h(t)$ (a small $E = E_{\min}$) assures a good approximation of the gain, real part, and imaginary part of $H(j\omega)$. Unfortunately, no similar assurance can be given for the phase approximation. A small $(H - H^*)$ leads to a small phase error $(\angle H - \angle H^*)$ only when $|H|$ or $|H^*|$ are much greater than $|H - H^*|$ Consequently, the phase error as $\omega \to \infty$ (where $|H|$ and $|H^*|$ generally go to zero) can be very large. As will be seen in later sections this deficiency of the minimum integral-square-error approximation can be corrected by obtaining a constrained approximation.

From these remarks it can be concluded that the integral-square-error criterion has quite reasonable approximation properties, particularly when constrained and weighted approximations are made.

Differential Analyzer Computation of the Approximation

Fig. 1 shows the scheme for computing a_n as given by (5). The approximating function $\phi_n(t)$ is realized as the impulse response of a differential analyzer circuit. The differential analyzer circuit is forced by an input $h(-t)$ (the prescribed impulse response reversed in time). From the superposition integral the output function of time is $\int_{-\infty}^{\infty} \phi_n(\tau) h(\tau - t) d\tau$. Thus at $t = 0$ the output is the desired coefficient a_n.[11]

To compute all a_n, all $\phi_n(t)$ must be realized as impulse responses of differential analyzer circuits. This will be described in the next section.

Once the a_n have been determined, $H^*(s)$ may be realized according to (7) with the differential analyzer circuits used for computation of the a_n. The arrangement is shown in Fig. 2. The summation box would

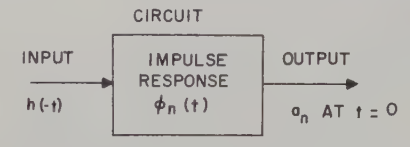


Fig. 1—Differential analyzer method for computation of a_n.

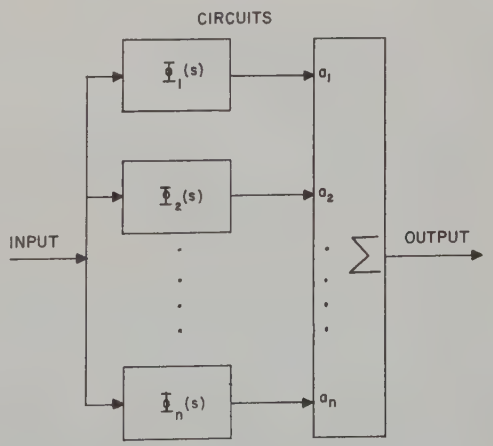


Fig. 2—Differential analyzer realization of $H^*(s)$.

involve coefficient potentiometers set to the a_n and two summing amplifiers (to take care of positive and negative a_n). In many cases the differential analyzer realiza-

[6] R. Courant and D. Hilbert, "Methods of Mathematical Physics," Interscience Publishers, Inc., New York, N. Y.; 1953.

[7] Y. W. Lee, "Synthesis of electrical networks by means of the Fourier transforms of Laguerre's functions," *J. Math. and Phys.*, vol. II, pp. 83–113; June, 1932.

[8] W. H. Kautz, "Transient synthesis in the time domain," IRE Trans. on Circuit Theory, vol. CT-1, pp. 29–39; September, 1954.

[9] E. G. Gilbert, "Linear System Approximation by Mean Square Error Minimization in the Time Domain," Ph.D. dissertation, University of Michigan, Ann Arbor, Mich.; 1956.

[10] Kautz, *op. cit.*, p. 35.

[11] The use of the convolution property to compute integrals is well known, *e.g.*, W. A. McCool, paper 14, Project Cyclone Symposium I, New York, N. Y.; March, 1951. A. G. Bose, "A Theory for the Experimental Determination of Optimum Nonlinear Systems," 1956 IRE Convention Record, pt. 4, pp. 21–31. W. H. Huggins, "Signal theory," IRE Trans. on Circuit Theory, vol. CT-3, pp. 210–216; December, 1956.

tion of H^* will suffice for the intended application of the approximate system. If a passive circuit realization of H^* is desired, it may be achieved by means of the usual techniques of network synthesis, subject to suitable restrictions on the gain and the singularities of $H^*(s)$.

If the realization in Fig. 2 is forced by an impulse its output is $h^*(t)$. Thus, the approximation error $(h-h^*)$ may be readily evaluated for all values of t. Such an evaluation is difficult when obtained from analytic computations.

A simple calculation shows how errors in computation of the a_n affect the integral square approximation error. Let $a_n+\Delta a_n$ be the computed coefficients and $E=E_{\min}+\Delta E$ the integral square error, increased from $E_{\min}$ by ΔE because of the coefficient errors Δa_n. Then

$$E = E_{\min} + \Delta E = \int_{-\infty}^{\infty}\left[\left(h - \sum_{n=1}^{N}(a_n + \Delta a_n)\phi_n\right)\right]^2 dt$$

$$= \int_{-\infty}^{\infty}\left[(h - h^*) - \sum_{n=1}^{N}\Delta a_n\phi_n\right] dt$$

$$= \int_{-\infty}^{\infty}(h - h^*)^2 dt - 2\sum_{n=1}^{N}\Delta a_n \int_{-\infty}^{\infty} h\phi_n dt$$

$$+ 2\sum_{n=1}^{N}\sum_{m=1}^{N}\Delta a_n a_m \int_{-\infty}^{\infty}\phi_n\phi_m dt$$

$$= +\sum_{n=1}^{N}\sum_{m=1}^{N}\Delta a_n \Delta a_m \int_{-\infty}^{\infty}\phi_m\phi_m dt$$

$$= E_{\min} - 2\sum_{n=1}^{N}\Delta a_n a_n + 2\sum_{n=1}^{N}\Delta a_n a_n + \sum_{n=1}^{N}\Delta a_n{}^2, \quad (12)$$

so that

$$\Delta E = \sum_{n=1}^{N}\Delta a_n{}^2. \quad (13)$$

When ΔE is small compared to $E_{\min}$ computer errors can be considered negligible. In the special case where the Δa_n are random with mean zero and variance $\sigma_a{}^2$, the average value of ΔE is simply $N\sigma_a{}^2$.

Orthonormal Approximating Functions

From (7) it is seen that the poles of $H^*(s)$ are the same as the poles of the $\Phi_n(s)$. Hence predetermination of the approximating functions amounts to predetermination of the poles of $H^*(s)$. Freedom in the choice of poles for $H^*(s)$ therefore requires that the formulation of the orthonormal functions allow poles at arbitrary values of s.

Such formulations are most easily expressed in the frequency domain. The set of functions $\Phi_1, \Phi_2, \cdots, \Phi_n$ considered here[12] have poles at $s=s_1, s_2, \cdots, s_n$ (the s_n must be in the left half-plane since $\int_{-\infty}^{\infty}\phi_n{}^2 dt=1$ requires the transfer function Φ_n to be stable). When s_n is real with $s_n=-\alpha_n$,

$$\Phi_n = \frac{(s+s_1)\cdots(s+s_{n-1})}{(s-s_1)\cdots(s-s_{n-1})}\,\frac{\sqrt{2\alpha_n}}{s+\alpha_n}. \quad (14)$$

When s_n and s_{n+1} are complex conjugates with real parts $-\alpha_n$ and magnitudes β_n,

$$\Phi_n = \frac{(s+s_1)\cdots(s+s_{n-1})}{(s-s_1)\cdots(s-s_{n-1})}\,\frac{2\beta_n\sqrt{\alpha_n}}{s^2+2\alpha_n s+\beta_n{}^2} \quad (15)$$

$$\Phi_{n+1} = \frac{(s+s_1)\cdots(s+s_{n-1})}{(s-s_1)\cdots(s-s_{n-1})}\,\frac{2\sqrt{\alpha_n}s}{s^2+2\alpha_n s+\beta_n{}^2}. \quad (16)$$

That the orthogonality conditions (3) hold for the above functions is shown in the frequency domain. The Parseval relation gives

$$\int_{-\infty}^{\infty}\phi_n\phi_m dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_n(j\omega)\Phi_m(-j\omega)d\omega = 0, \quad n\neq m$$

$$= 1, \quad n=m. \quad (17)$$

Since $\Phi_n(j\omega)\Phi_m(-j\omega)$ goes to zero more rapidly than $1/j\omega$ as $\omega\to\infty$, (17) can be expressed as (the contour C is shown in Fig. 3 and encircles the entire right half-plane)

$$\frac{1}{2\pi j}\int_C \Phi_n(s)\Phi_m(-s)ds$$

$$= \sum \text{residues of } \Phi_n(s)\Phi_m(-s) \text{ in right half plane}$$

$$= 0, \quad n\neq m$$

$$= 1, \quad n=m. \quad (18)$$

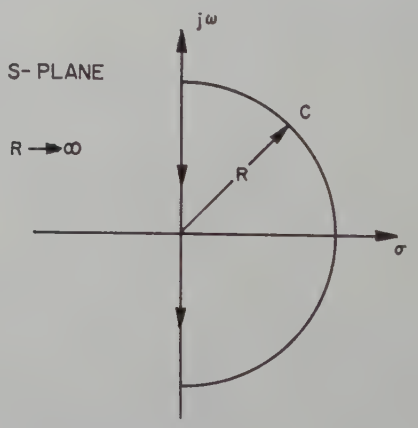


Fig. 3—Contour for integral in (18).

Using the functions (14) in $\Phi_n(s)\Phi_m(-s)$ and defining n to be the greater of the two indices gives

$$\Phi_n(s)\Phi_m(-s)$$

$$= \frac{(s+s_1)\cdots(s+s_{n-1})}{(s-s_1)\cdots(s-s_{n-1})}\,\frac{\sqrt{2\alpha_n}}{s+\alpha_n}$$

$$\cdot\frac{(-s+s_1)\cdots(-s+s_{m-1})}{(-s-s_1)\cdots(-s-s_{m-1})}\,\frac{\sqrt{2\alpha_m}}{-s+\alpha_m}$$

$$= \frac{(s+s_m)\cdots(s+s_{n-1})}{(s-s_m)\cdots(s-s_{n-1})}\,\frac{\sqrt{2\alpha_n}}{s+\alpha_n}\,\frac{\sqrt{2\alpha_n}}{-s+\alpha_m}. \quad (19)$$

[12] Other similar formulations are possible. Kautz or Gilbert, *op. cit.*

But $\alpha_n = -s_m$ so $\Phi_n(s)\Phi_m(-s)$ has all of its poles in the left half-plane. Thus there are no residues of $\Phi_n(s)\Phi_m(-s)$ in the right half-plane and (18) is satisfied for $n \neq m$. For $n = m$ it is easily shown that (18) is also satisfied. Similar results ensue when any combination of the functions (14), (15), and (16) is substituted in (18).

Since Φ_1 has a single pole at s_1, Φ_2 has two poles at s_1 and s_2, $\cdots$, and Φ_n has n poles at s_1, s_2, $\cdots$, s_n, differential analyzer realization of the Φ_n is most easily achieved by a cascade connection of sections as shown in Fig. 4. $F_0(s)$ is the transform of input; the $F_0(s)\Phi_n(s)$ are the transforms of the outputs, and

$$F_n(s) = \frac{(s + s_1) \cdots (s + s_n)}{(s - s_1) \cdots (s - s_n)} F_0(s).$$

Operational amplifier circuits for the sections in Fig. 4 are shown in Fig. 5. Except for the normalization factors

$$\frac{\pm 1}{2\sqrt{\alpha_n}} \quad \text{and} \quad \frac{-1}{\sqrt{2\alpha_n}},$$

these sections may be fitted directly into Fig. 4. For the complex pole section $F_n(s)$ does not appear since $f_n(t)$ is a complex function of time. Note that the pole positions of the Φ_n are defined by the potentiometer settings α_n and β_n. Thus the positions may be shifted by a simple computer adjustment without any analytic calculations.

The ease with which the pole positions of the Φ_n are varied and $E_{\min}$ is computed by (6) allows many different combinations of pole positions to be tried for a given $h(t)$. The means that a good set of pole positions can be obtained by trial and error and need not depend on complicated procedures or the intuition of the designer.

Constrained and Weighted Approximations

It has been pointed out that constrained and weighted approximations correct certain deficiencies of the integral-square-error criterion. This section describes simple modifications of the methods already presented to achieve such approximations.[13]

A constrained approximation is one in which the coefficients in the series h^* are functionally related so that a property of h^* is specified. Mathematically, a constraint may be expressed by

$$k = K(h^*(t)) = K\left(\sum_{n=1}^{N} b_n \phi_n(t)\right), \tag{20}$$

where K is a functional describing a property of h^*, k is the specified value of the property, and

$$h^* = \sum_{n=1}^{N} b_n \phi_n$$

[13] The procedures of this section have been described in slightly different form by Gilbert, *op. cit.* and W. H. Kautz, "Network Synthesis for Specified Transient Response," Tech. Rep. No. 209, M.I.T. Res. Lab. of Electronics, Cambridge, Mass.; April, 1952.

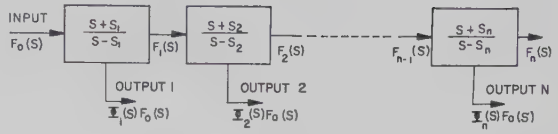
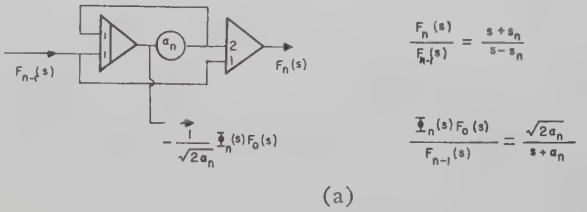
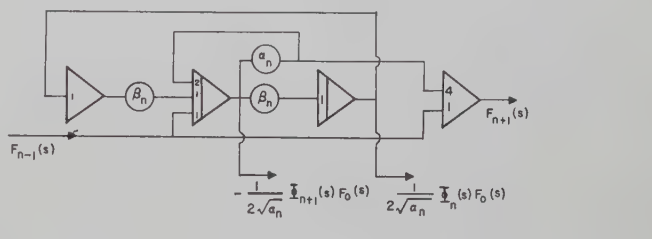


Fig. 4—Differential analyzer realization of $\Phi_n(s)$.



(a)



(b)

Fig. 5—Operational amplifier circuits for sections shown in Fig. 4.

is the constrained approximation. Examples of (20) are:

1) The area under h^* equals unity (this is equivalent to saying $H^*(0) = 1$)

$$1 = \int_{-\infty}^{\infty} h^* dt = \sum_{n=1}^{N} b_n \int_{-\infty}^{\infty} \phi_n dt.$$

2) $h^* = 0$ at $t = 0$ (this equivalent to saying that $H^* \to (\text{const}/s^2)$ as $s \to \infty$, *i.e.*, $H^*(j\omega)$ has phase shift $\pm 180°$ at $\omega = \infty$)

$$0 = h^*(0) = \sum_{n=1}^{N} b_n \phi_n(0).$$

3) The integral square value of h^* equals unity

$$1 = \int_{-\infty}^{\infty} h^{*2} dt = \sum_{n=1}^{N} \sum_{n=1}^{N} b_n b_m \int_{-\infty}^{\infty} \phi_n \phi_m dt = \sum_{n=1}^{\infty} b_n{}^2.$$

Conditions 1 and 2 are linear in the b_n; condition 3 is nonlinear. Because of computational difficulties, nonlinear constraints are not considered here.

The constrained approximation is obtained by the method of Lagrange multipliers.[14] Since $k - K(h^*) = 0$, minimizing

$$E_C = \int_{-\infty}^{\infty} (h - h^*)^2 dt - \lambda(k - K(h^*)), \tag{21}$$

[14] Courant and Hilbert, *op. cit.*, p. 165.

with respect to the b_n and λ is equivalent to minimizing $\int_{-\infty}^{\infty}(h-h^*)^2dt$ subject to (20). Setting $\partial E_C/\partial b_n=0$, $n=1, \cdots, N$ and $\partial E_C/\partial\lambda=0$ gives

$$b_n = a_n + \tfrac{1}{2}\lambda K(\phi_n), \qquad n = 1, \cdots N \tag{22}$$

and

$$k = \sum_{n=1}^{N} b_n K(\phi_n) = \sum_{n=1}^{N} (a_n + \tfrac{1}{2}\lambda K(\phi_n))K(\phi_n). \tag{23}$$

The a_n are found as before from (5); λ is determined from (23); (22) gives the constrained approximation coefficients b_n. The same general procedure holds when more than one constraint equation, (20), is given except more Lagrange multipliers and constraint conditions are introduced in (21).

Since $b_n \neq a_n$, $E_{C\min} > E_{\min}$. By letting $\Delta a_n = b_n - a_n$, the deviation of the b_n from the unconstrained approximation coefficients, (13) may be applied to give

$$E_{C\,\min} - E_{\min} = \Delta E = \sum_{n=1}^{N} (b_n - a_n)^2. \tag{24}$$

Here ΔE is the integral-square-error penalty incurred for satisfaction of the constraint equation.

Minimization of the weighted integral-square-error

$$E_W = \int_{-\infty}^{\infty} (h - h^*)^2 W(t)dt \tag{25}$$

results in better control of the approximation error. Where $W(t)$ is made large, $(h-h^*)$ tends to be small. The most straightforward way of minimizing E_W is to use functions $\phi_n(t)$ orthonormal with respect to the weight factor $W(t)$. Unfortunately, no simple frequency domain representation [*e.g.*, (14)–(16)] of such functions is possible.

To proceed using the functions given by (14), (15), and (16), E_W is written as the unweighted integral-square-error

$$E_W = \int_{-\infty}^{\infty} (W^{1/2}h - W^{1/2}h^*)^2 dt. \tag{26}$$

Then $W^{1/2}h$ may be approximated by

$$W^{1/2}h^* = \sum_{n=1}^{N} a_n\phi_n, \tag{27}$$

where $W^{1/2}h$ assumes the role of h in previous developments. From (5)

$$a_n = \int_{-\infty}^{\infty} W^{1/2}h\phi_n dt. \tag{28}$$

The weighted approximation is

$$h^* = W^{-1/2} \sum_{n=1}^{N} a_n\phi_n. \tag{29}$$

Since h^* is realized by lumped element systems, it must be a sum of exponential functions. According to (29) this means that $W^{-1/2}$ must also be a sum of exponential functions, say

$$W^{-1/2} = \sum_{k=1}^{K} w_k\epsilon^{\bar{s}_k t}. \tag{30}$$

Then

$$h^* = \sum_{k=1}^{K} w_k\epsilon^{\bar{s}_k t} \sum_{n=1}^{N} a_n\phi_n(t), \tag{31}$$

and

$$H^* = \sum_{k=1}^{K} w_k \sum_{n=1}^{N} a_n\Phi_n(s - \bar{s}_k). \tag{32}$$

Two disadvantages of the method are clear: 1) The class of weight factors is limited. 2) The approximation requires KN rather than N approximating functions. By making K sufficiently large, disadvantage 1) can be overcome in most practical problems. This makes H^* very complicated since KN then becomes large. In many cases, $K=1$ or 2 gives a satisfactory compromise of the two disadvantages. A final restriction must be observed. $W^{-1/2}$ must not go to zero in such a way as to cause (28) to become unbounded.

Evaluation of (28) is obtained as before except that $h(-t)W^{1/2}(-t)$ replaces $h(-t)$. Realization of the functions $\Phi_n(s-\bar{s}_k)$ is easily accomplished when the $\bar{s}_k$ are real by the circuit substitutions shown in Fig. 6.

Examples

The first example is an unweighted, unconstrained approximation of a fourth power pulse

$$\begin{aligned} h(t) &= \frac{t^4}{256} - \frac{t^3}{16} + \frac{t^4}{4}, \qquad 0 \leq t \leq 8 \\ &= 0, \qquad t < 0, \quad t > 8, \end{aligned} \tag{33}$$

which is symmetric about $t=4$, has a maximum $h(4)=1$, and has zero value and slope at $t=0$ and $t=8$. Six approximating functions are used and all $s_n = -1(\alpha_n = 1)$.[15] Since all s_n is real (14) defines the $\Phi_n(s)$. The only consideration which entered in the choice of the s_n was to make the spread of the widest function $(\phi_6(t))$ approximate the spread of $h(t)$. Fig. 7, obtained from the differential analyzer, shows that the approximation is quite good.

The coefficients a_n were computed analytically and are exact within the number of significant figures given; the Δa_n are errors due to differential analyzer computation of the coefficients.[16] The relative-integral-square error, defined by

$$e_{\min} = \frac{E_{\min}}{\int_{-\infty}^{\infty} h^2 dt}, \tag{34}$$

[15] Except for multiplicative constants these $\phi_n(t)$ are the Laguerre functions.

[16] Static errors of machine components were on the order of 0.1 per cent.

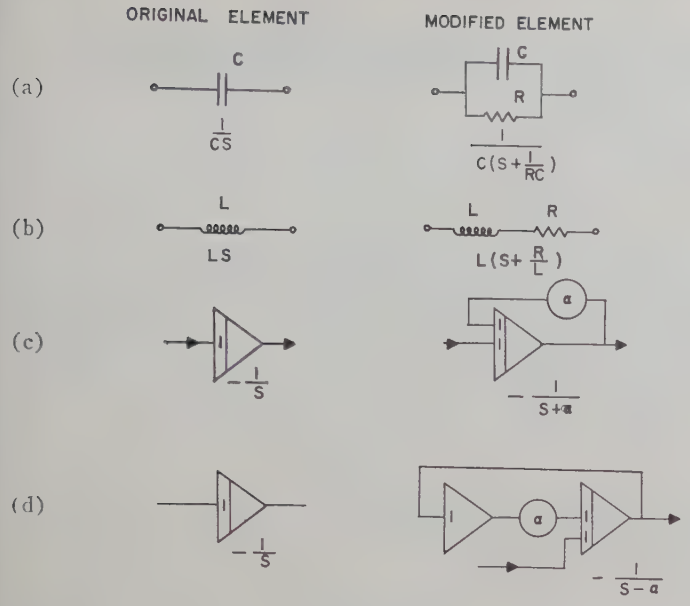


Fig. 6—Circuit modifications for changing s into $(s-\bar{s}_k)$.

is 0.00651. Δe, the relative integral-square-error due to the Δa_n, is computed by (13) and is only 6×10^{-6}. Thus differential analyzer errors are completely negligible. They will not be considered in what follows.

Since h^* and h are not the same at $t=0$, H^* and H do not fall off at the same rate as $s\to\infty$. This results in a large phase error as $j\omega\to j\infty$. To eliminate this approximation discrepancy h^* must be constrained so that

$$h^*(0) = h(0) = 0 \text{ and } \left.\frac{dh^*}{dt}\right|_{t=0} = \left.\frac{dh}{dt}\right|_{t=0} = 0.$$

Fig. 8 shows an approximation where only $h^*(0)=0$. The a_n are the $a_n+\Delta a_n$ given in Fig. 7; (23) gives $\lambda=-0.028$; (22) yields the b_n. Leniency of the imposed constraint is indicated by the small increase in the relative integral-square-error from $e_{\min}=0.00651$ to $e_{C\min}=0.00687$.

As a final example the trapezoid pulse

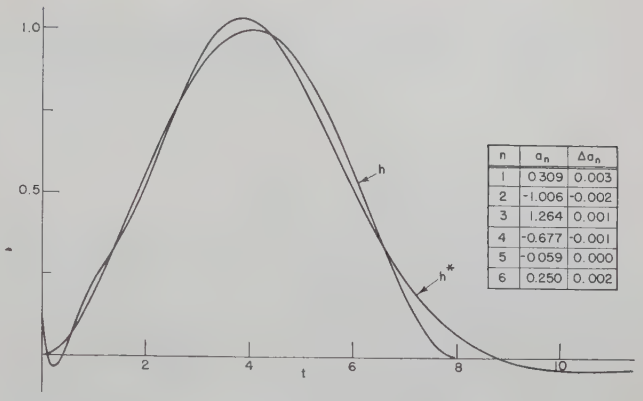
$$\begin{aligned} h(t) &= 0, && t<0, \quad t>4,\\ &= 1, && 0\le t<2,\\ &= 2-\tfrac{1}{2}t, && 2\le t\le 4, \end{aligned} \tag{35}$$

is approximated. Five approximating functions are employed, with poles taken rather arbitrarily to be on the unit circle at

$$s=-1,\ -\tfrac{2}{3}\pm j\sqrt{1-(\tfrac{2}{3})^2},\ -\tfrac{1}{3}\pm j\sqrt{1-(\tfrac{1}{3})^2}.$$

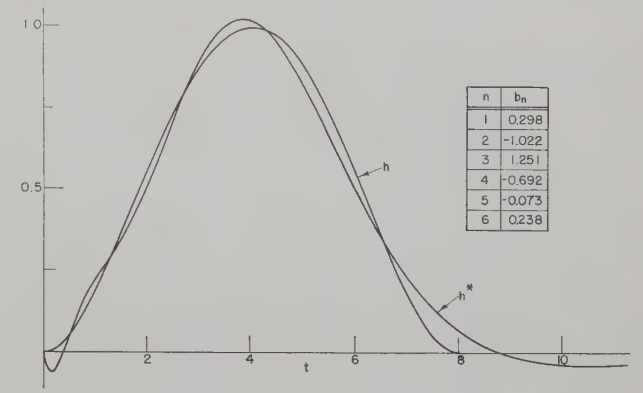
Differential analyzer realization of the $\Phi_n(s)$ requires one (a) section and two (b) sections of the type shown in Fig. 5. In the setup used the sections were ordered in the same way as the poles are given above, although any order would, of course, work.

For $W(t)=1$, h^* and the a_n are given in Fig. 9. The relative integral-square-error is $e_{\min}=0.0098$. Fig. 9 also shows an approximation for $W(t)=e^{0.5t}$. Since $K=1$,



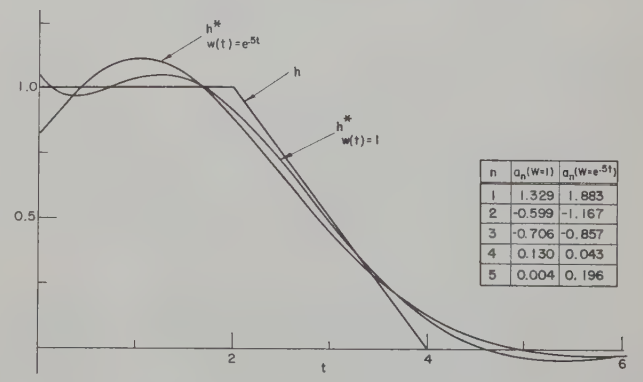
n	a_n	Δa_n
1	0.309	0.003
2	-1.006	-0.002
3	1.264	0.001
4	-0.677	-0.001
5	-0.059	0.000
6	0.250	0.002

Fig. 7—Unconstrained approximation of fourth power pulse.



n	b_n
1	0.298
2	-1.022
3	1.251
4	-0.692
5	-0.073
6	0.238

Fig. 8—Constrained approximation of fourth power pulse.



n	$a_n(W=1)$	$a_n(W=e^{5t})$
1	1.329	1.883
2	-0.599	-1.167
3	-0.706	-0.857
4	0.130	0.043
5	0.004	0.196

Fig. 9—Weighted and unweighted approximation of trapezoid pulse.

$W^{-1/2}(t)=e^{-0.25t}$ and (32) becomes

$$H^* = \sum_{n=1}^{5} a_n\Phi_n(s+0.25). \tag{36}$$

In order to have the same poles in H^* as in the approximation for $W(t)=1$, the s_n are:

$$s_1=-\tfrac{3}{4};\ s_2, s_3=-\frac{5}{12}\pm j\sqrt{1-(\tfrac{2}{3})^2};\ s_4, s_5=-\frac{1}{12}\pm j\sqrt{1-(\tfrac{1}{3})^2}.$$

The relative weighted integral-square-error is $e_{W\min}=0.0220$. As expected, the weighting factor causes the error to be decreased for large t and increased for small t.

Generalized Integration on the Analog Computer*

GEORGE A. BEKEY†

Summary—One of the major limitations of the electronic analog computer is its inability to perform directly an integration with respect to a dependent variable. This paper reviews the usual methods of overcoming this limitation, describes the results on an attempt to use Padé time-delay units in generalized integration, and presents the development of a new analog integrator based on a simple numerical integration formula. The integrator can be instrumented using standard analog computer components. The performance of the device is illustrated with several examples.

Introduction

ONE of the major limitations of the electronic analog computer is its inability to perform directly an integration with respect to a dependent variable. The mechanical differential analyzer,[1] which does not suffer from this limitation, is thus a more versatile machine. It was displaced by the electronic differential analyzer primarily because of its large size, slow speed of operation, and the excessive time required to place a problem on the machine. However, many of those who had their early training on a mechanical analog computer continue to think with nostalgia of the days when one could integrate any variable in the computer with respect to any other variable. Perhaps the fact that the electronic analog computer lacks the versatility of the mechanical machine, has led to its increasingly restricted applications to only certain types of time-varying problems. The gradual disappearance of the term "differential analyzer" may thus be due to the fact that the electronic analog computer is not a general purpose differential equation solver, in contrast to the mechanical or digital differential analyzer.

The ability to perform integration with respect to an arbitrary variable lends new flexibility to problem formulation. Sometimes differentiating an equation leads to considerable simplification, but may introduce differentials of dependent variables if the equation is nonlinear or has time-varying coefficients. The formulation of certain problems in polar coordinates requires simultaneous integration with respect to r and θ. The generation of certain analytic functions of dependent variables and the transformation of polar to rectangular coordinates are further examples of operations which may be facilitated by performing integration with respect to dependent variables.

* Manuscript received by the PGEC, November 13, 1958; revised manuscript received February 19, 1959. This paper was presented at the National Simulation Conference, Dallas, Tex., October 23–25, 1958.

† Space Tech. Labs., Inc., Los Angeles, Calif.

[1] V. Bush and S. H. Caldwell, "A new type of differential analyzer," *J. Franklin Inst.*, vol. 240, pp. 255–326; 1945.

Mechanical integration is performed by means of several variations of the Kelvin disc integrator, which is shown schematically in Fig. 1. If the turntable is turned through an angle Δx, the corresponding angular displacement of the small wheel will depend on its radius and distance from the turntable axis, *i.e.*,

$$\Delta z = ky\Delta x. \tag{1}$$

Since this relationship holds for infinitesimal motions (provided the small wheel does not slip), we have

$$z = k\int ydx \tag{2}$$

where y and x can be any two variables. Neither is restricted to being monotonic or of a single polarity. The operation of (2) can be performed on the electronic analog computer by making use of

$$z = \int_{x_1}^{x_2} ydx = \int_{t_1}^{t_2} y\left(\frac{dx}{dt}\right)dt \tag{3}$$

where it is assumed that $x(t)$ is a monotonic function so that the change of limits indicated above can be performed. If $x(t)$ is not monotonic, t becomes a multiple valued function of x, and (3) may be valid only over certain intervals of t where it is single valued. Thus, the evaluation of the general integral requires differentiation, multiplication, and time-integration as indicated in Fig. 2. Alternative methods of generalized integration thus depend on the choice of method for approximate differentiation. The first part of this paper is concerned with several methods of differentiation, including time-delay techniques and their use in performing generalized integration. The second portion of the paper deals with the development of an analog integrator based on a simple numerical integration formula.

Methods of Differentiation

Of the components involved in performing integration by the method of (3), the differentiator poses the greatest problems. Its frequency response must be restricted in order to keep noise within acceptable amplitude limits without introducing excessive phase-shift. The over-all accuracy of the integration will then be a composite of the accuracy of differentiation, multiplication, and final time-integration. Two types of approximate differentiation circuits are shown in Fig. 3. The output of the circuit of Fig. 3(a), assuming amplifier gain to be infinite, is given by

$$e_0 = -\frac{R_2Cs}{(R_1C)s + 1}e_i. \tag{4}$$

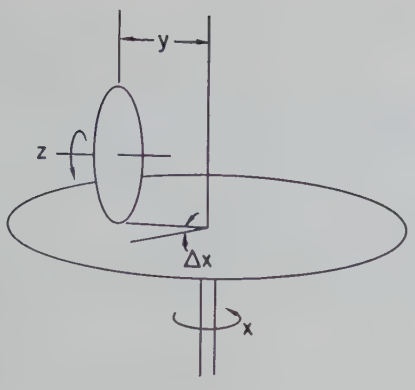


Fig. 1—A Kelvin disc integrator.

If the time constant $(R_1C) \ll R_2C$, this reduces to

$$e_0 \cong -(R_2C)se_i. \tag{5}$$

In practice, R_1 is often adjusted for optimum accuracy within the limits of acceptable noise level at the output.

Fig. 3(b) represents a method of implicit differentiation. The high-gain amplifier enforces the relationship

$$e_i + e_2 = -\frac{e_0}{A} \cong 0 \tag{6}$$

where

$$e_2 = \frac{e_0}{s} \tag{7}$$

and, therefore,

$$e_0 \cong -se_i. \tag{8}$$

A similar circuit can be constructed using a two-amplifier positive-feedback, unity-gain loop instead of the high-gain amplifier. The filter condenser C_s is required to limit the noise level of the output and to insure stability.

While the circuits shown in Fig. 3 are satisfactory for use at low frequencies, they are not usable with wide bandwidth signals because of the nature of the process of differentiation and the noise which is present in analog signals. In an effort to circumvent the computer noise problem, a circuit to obtain the average derivative by use of a time delay, was investigated.

Differentiation with a Time Delay

The availability of packaged time-delay units makes possible a different method of differentiation for use with a generalized integrator. The method is based on

$$\frac{dx}{dt} = \lim_{\tau \to 0} \frac{x(t) - x(t-\tau)}{\tau} \tag{9}$$

as indicated in block diagram form in Fig. 4. The time-

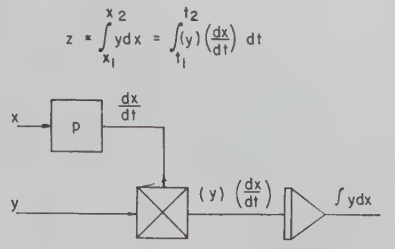


Fig. 2—Conventional method of generalized integration.

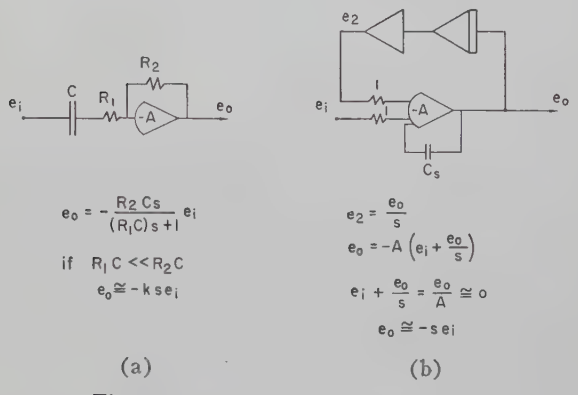


Fig. 3—Two methods of differentiation.

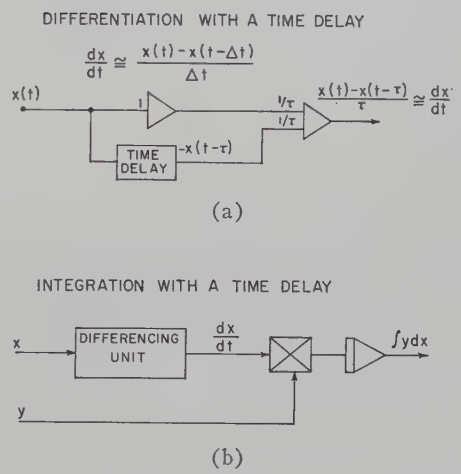


Fig. 4—Use of time-delay unit in differentiation and generalized integration.

delay unit provides the voltage $-x(t-\tau)$, which is summed with $x(t)$ and multiplied by $1/\tau$ to approximate the derivative.

The time-delay unit used was a fourth-order Padé approximation.[2,3] This approximation to the delay function $f(t-\tau)$ is based on the Laplace transform relationship

$$L[f(t-\tau)] = F(s)e^{-\tau s}. \tag{10}$$

[2] C. H. Single, "Time Delay Circuits for Analog Computers," Berkeley division of Beckman Instruments, Inc., Richmond, Calif.; March 10, 1956.

[3] W. J. Cunningham, "Time delay networks for an analog computer," IRE Trans. on Electronic Computers, vol. EC-3, pp. 16–18; December, 1954.

The exponential $e^{-\tau s}$ is approximated by a rational function in which the numerator and denominator are polynomials in s. If the numerator and denominator are of the same degree, the function exhibits constant amplitude at all frequencies and linear phase shift to some maximum frequency. The fourth-order approximation is given by

$$e^{-\tau s} \cong \frac{(\tau s)^4 - 20(\tau s)^3 + 180(\tau s)^2 - 840\tau s + 1680}{(\tau s)^4 + 20(\tau s)^3 + 180(\tau s)^2 + 840\tau s + 1680}. \quad (11)$$

The bandwidth of this unit is approximately $f = 1.2/\tau$ cps.

Accuracy and Frequency Response

The accuracy of the unit is clearly a function of the size of delay interval τ, the characteristics of the delay unit, and any gain mismatch at the inputs to the summing amplifier. The calibration of the delay unit used in this study decreased in precision rapidly for very small delay times; and yet short delays were necessary in order to achieve sufficient accuracy over a reasonable bandwidth. An error analysis of the differentiator for a sinusoidal input is given in the Appendix and shows that values of the product $(\omega\tau)$ as low as 0.02 are required to obtain an accuracy of 1 per cent. With Padé delays, about 0.01 second was the minimum delay obtainable, thus restricting the maximum frequency to 2 radians per second.

Noise Level

The output of the time delay unit utilized is the output of an operational amplifier, and its noise level is essentially independent of the delay setting. The noise at the output of the differentiator is thus inversely proportional to the selected delay interval τ and directly proportional to the desired bandwidth. The output noise will consist of amplifier noise multiplied by $1/\tau$, or input signal noise times $1/\tau$, whichever is greater.

Examples

Fig. 5 shows the error obtained in differentiating the function $x = A \cos \omega t$. It is shown in the Appendix that the expected error is

$$\epsilon = A\left[\sin \omega t\left(\frac{\sin \omega\tau}{\omega\tau} - 1\right) - \cos \omega t\left(\frac{1 - \cos \omega\tau}{\omega\tau}\right)\right]. \quad (12)$$

The error thus includes an in-phase and an out-of-phase component, the magnitudes of which depend on the factor $(\omega\tau)$. With $\omega = 1$ and a delay setting $\tau = 0.01$ second the error amplitude is approximately 1 per cent of the output amplitude A.

The need for extremely critical calibration of the delay time is illustrated in Fig. 6, which shows error curves from the differentiation of $x = A \sin \omega t$. The summing amplifier input gains were based on a delay interval $\tau = 0.05$ second, *i.e.*, a gain of 20 for $\omega = 1$. Thus, delay-setting errors were magnified by a large factor. Fig. 6 shows error curves for dial settings ranging from 0.035–0.045 second. The minimum error curve (dial setting $= 0.040$ second) checks with the predicted error magnitude. Other types of delay units with more precisely adjustable delay times would be preferable for high accuracy.

The use of the time-delay differentiator in a complete generalized integration circuit is shown in Fig. 7. The operation performed was

$$f = 1/2 \int_0^x y dx$$

$$y = \frac{t^2}{2}$$

$$x = t. \quad (13)$$

The result of this simple integration is

$$f = 1/2 \int_0^t \left(\frac{t^2}{2}\right) dt = \frac{t^3}{12} \quad (14)$$

and the error expected will be

$$\epsilon = \frac{t^3}{12} - 1/4 \int \frac{t}{\tau}\left[\frac{t^2}{2} - \frac{(t-\tau)^2}{2}\right] dt \quad (15)$$

$$\epsilon = \frac{\tau}{8} t^3. \quad (16)$$

The error curve is thus also a function of t^3. The output of the integration and the error curve are shown in Fig. 7. For this problem the error after 10 seconds of integration was less than 0.5 per cent of the value of integral.

In view of the fact that a 0.02 second time delay is required for a 1 per cent differentiation of a 1 radian per second sinewave, it is apparent that the use of the time-delay approximation for differentiation is not practical for use in any wide-band application. It is also evident that all generalized integration methods which depend on differentiation are subject to serious limitations in accuracy, noise, and frequency response. The following section represents an alternative scheme for generalized integrations which does not depend directly on differentiation of the integrand.

A New Generalized Integration Method

The basis of the method here presented is the computer mechanization of a simple numerical integration formula. Consider the sketch of Fig. 8. The area under the curve between appropriate limits represents the desired integral. If we divide the area into a series of vertical strips of equal width Δx, and let the values of the ordinate at the center of each strip be represented by y_i, we can write

$$F = \int_{x_1}^{x_n} y dx = \lim_{\Delta x \to 0} \sum_{i=1}^{n} y_i \Delta x \quad (17)$$

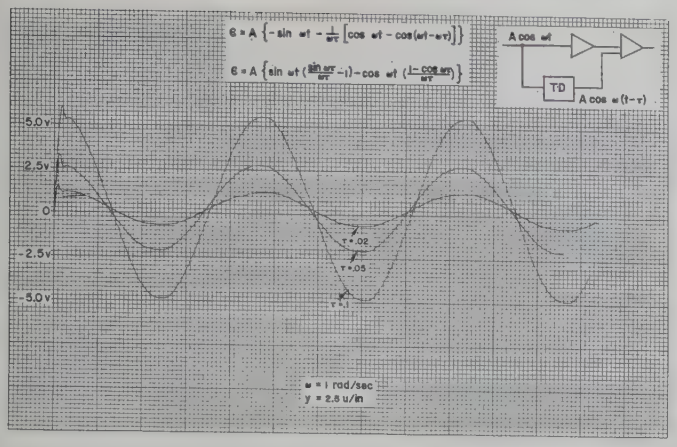


Fig. 5—Error curves resulting from the differentiation of $x=A\cos \omega t$ with a time-delay circuit for several values of delay.

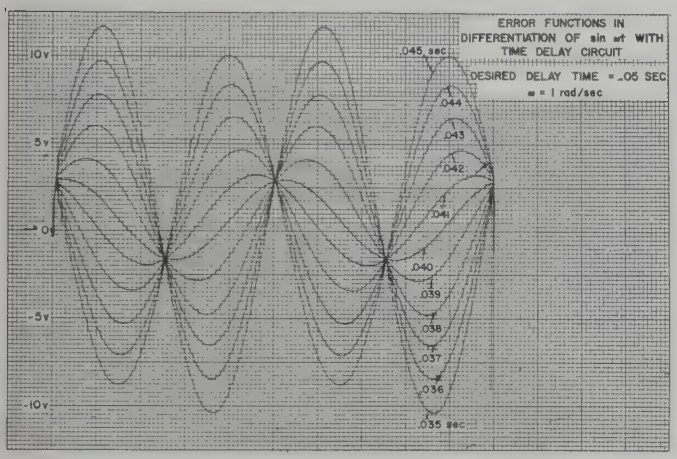


Fig. 6—Error curves of a time-delay differentiator as a function of delay setting.

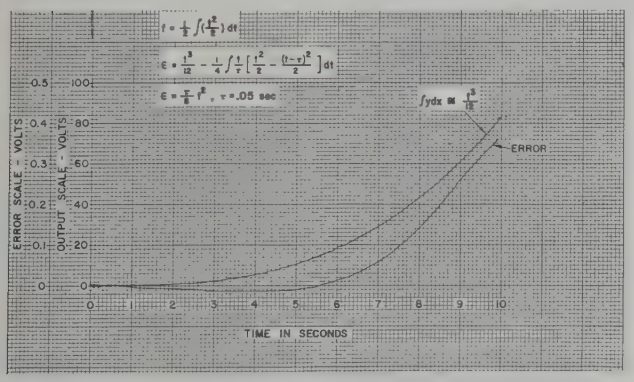
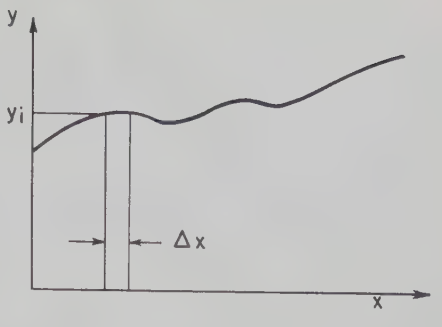


Fig. 7—Result of generalized integration using a time delay for a simple problem.

and, if Δx is small,

$$\therefore\ F \cong (\Delta x)\sum_{i=1}^{n} y_i. \qquad (18)$$

Thus, the evaluation of (17) requires the summation of a series of values of y, which are obtained by sampling at constant intervals in the variable of integration x.



$$F = \int_{x_1}^{x_2} y\,dx \cong \sum_{x=x_1}^{x_2} y_i \Delta x = (\Delta x)\sum_{x=x_1}^{x_2} y_i$$

Fig. 8—Approximate integration by summation of areas.

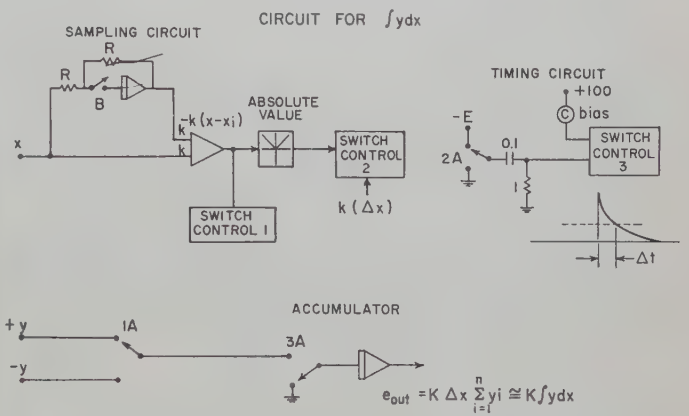


Fig. 9—Schematic diagram of new generalized integrator.

Mechanization of the Integrator

The mechanization of (18) was performed using standard analog computer components to evaluate performance and determine needed specifications. The circuit is shown in Fig. 9. The behavior of this circuit can be clarified by considering the following points:

1) Integrator 1 is connected as a "sample-and-hold" device,[4,5] so that its output represents the value of x at the last sample, x_i.
2) Since the integrator is not restricted to monotonically increasing functions of time, the quantity $(x-x_i)$ may take on either positive or negative values. Its sign is used to determine whether $+y$ or $-y$ is sampled.
3) Sampling of x and y is performed when $|x-x_i| = \Delta x$, the preset width of the sampling interval.
4) The sampled values of y are summed in an integrator which behaves as an accumulator by

[4] L. B. Wadel, "Analysis of combined sampled and continuous data systems on an analog computer," 1955 IRE National Convention Record, vol. 3, pt. 4, pp. 3–7.

[5] H. Chestnut, A. Dabul, and D. Leiby, "Analog computer study of sampled data systems," *Proc. Computers in Control Systems Conf.*, pp. 71–76; October, 1957.

integrating the input voltage for a fixed period of time (Δt). If Δt is sufficiently small, we can assume y will remain constant during this period, so that the output of the accumulator will consist of a staircase of integrated inputs.

5) A timing circuit, consisting of a clipped pulse of controllable width, determines the integrating time (Δt).

The output of the integrator at the nth sampling time is given by

$$e_{0n} = e_{0(n+1)} + \int_{t_n}^{t_n+\Delta t} y_n dt \tag{19}$$

where $e_{0(n-1)}$ is the output of the integrator after the $(n-1)$st sampling pulse. Therefore,

$$e_{0n} = e_{0(n-1)} + y_n \Delta t. \tag{20}$$

Since each sample contributes a voltage $y_i \Delta t$,

$$e_{0n} = (\Delta t) \sum_{i=1}^{n} y_i. \tag{21}$$

Thus, the value of the desired integral can be obtained from the integrator output by noting that since

$$F = \int_{x_1}^{x_n} y dx \cong (\Delta x) \sum_{i=1}^{n} y_i, \tag{22}$$

we can write

$$F \cong \frac{(\Delta x)}{(\Delta t)} e_{0n}. \tag{23}$$

The output of the final integrator is therefore proportional to value of the generalized integral.

Equipment Considerations

Relay amplifiers were used for the switching operations indicated in Fig. 9. The relays used have throw times of the order of 1–2 msec. Since considerably faster switching rates would be required in a commercial device of reasonable bandwidth, electronic switches will be necessary.

The integration period Δt must be long compared to the switch throw time to obtain repeatable performance. With the relays used in this study it was not possible to use integration periods shorter than 30 msec.

The sampling circuit is simply a first-order lag which charges up to the input voltage with a time constant RC, where C is the integrating condenser. The timing circuit then not only applies the y-voltage to the accumulator for a period Δt, but closes the sampling circuit switch for the same length of time. Using $R=15$K and $C=0.1$ μf, the time constant of the sampling circuit was 1.5 msec, and the total sampling period of 30 msec was equivalent to 20 time constants.

The sampling interval (Δx) is in volts rather than seconds, and must be selected with the consideration that at the highest frequency of interest in the variable x, the time between two samples in x will be sufficiently long compared to the integration interval Δt. Waveforms in the circuit are shown in Fig. 10.

Specifications

Experimentation has shown that the smallest time constant achievable with the unmodified Electronic Associates amplifier, when used as a first-order lag, is 0.05 msec ($R=10$K, $C=0.005$ μf). The rise time of the amplifier is seen if shorter time constants are attempted. Thus, the sampling pulse width (switch closure time) should be about 0.5 msec, to allow 10 time constants for the circuit to "charge up" to within 0.01 per cent of the input. If $\Delta t=0.5$ msec, the samples in x should be spaced sufficiently widely to allow Δt to be smaller than the maximum time between x samples. If we arbitrarily choose this time to be 1 msec $=(2\Delta t)$ this determines the maximum possible slope of x, *i.e.*,

$$\left(\frac{dx}{dt}\right)_{\max} = \frac{\Delta x}{2\Delta t}. \tag{24}$$

Assume that Δx is chosen as 0.1 volt. Then the maximum slope of x will be restricted to

$$\left(\frac{dx}{dt}\right)_{\max} = \frac{0.1 \text{ volt}}{0.001 \text{ sec}} = 100 \text{ volts per second}. \tag{25}$$

This is a rather restrictive slope, since it corresponds to the maximum slope of a 100-volt amplitude sine wave of frequency $\omega=1$ radian per second. Further improvement in the sampling circuit time constant will be sought in the future.

Electronic switches, such as a four-diode gate,[6] can be used at considerably higher switching speeds than discussed above, and appear to present no serious obstacle. Accuracies of the order of 0.01 per cent for low frequencies (below 1 cps) appear reasonable with this device. The accuracy will decrease progressively for higher frequencies in either x or y.

Results

The circuit was tested by obtaining the function $F=\int y dx$ for various types of input functions. Fig. 11 shows the result of setting $y=$constant, $x=kt$. Then

$$F = \int c d(kt) = ckt. \tag{26}$$

The stepwise nature of the output is clearly visible in this output curve, which was obtained on an x-y plotter.

[6] N. Diamantides, "Electronic switch for analog computer simulation," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-5, pp. 197–202; December, 1956.

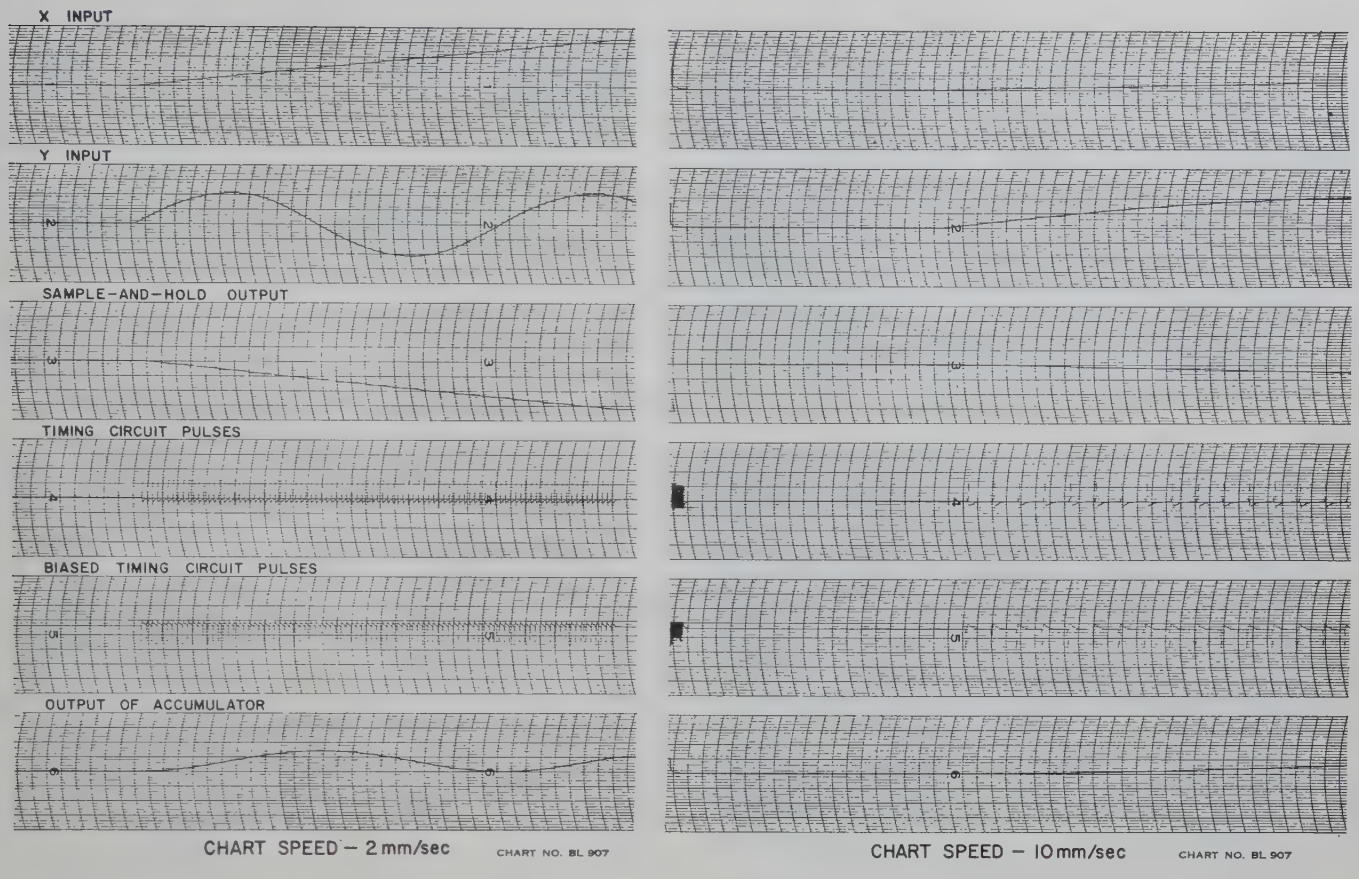


Fig. 10—Output waveforms in the generalized integrator.

Setting $x = y = kt$ produces the result of Fig. 12.

$$F = \int ktd(kt) = \frac{k^2t^2}{2}. \tag{27}$$

As a first experiment with a nonmonotonic x-input, the following integration was performed:

$$F = \int cd(k \sin \omega t) = ck \sin \omega t. \tag{28}$$

The result, for two sampling widths Δx, is shown in Fig. 13. Note that since the circuit does not sample when x is not changing (*i.e.*, when $dx/dt = 0$), the output sine waves have flat peaks.

Finally, as a simple test of the circuit for two varying inputs, the function evaluated was

$$F = \int \sin \omega td(\sin \omega t) \tag{29}$$

$$F = \frac{\sin^2 \omega t}{2} = 1/4(1 - \cos 2\omega t). \tag{30}$$

The output waveform is sinusoidal, but of double the frequency of the inputs, as shown in Fig. 14. Once again, the influence of dx/dt upon the switching rate is clearly evident.

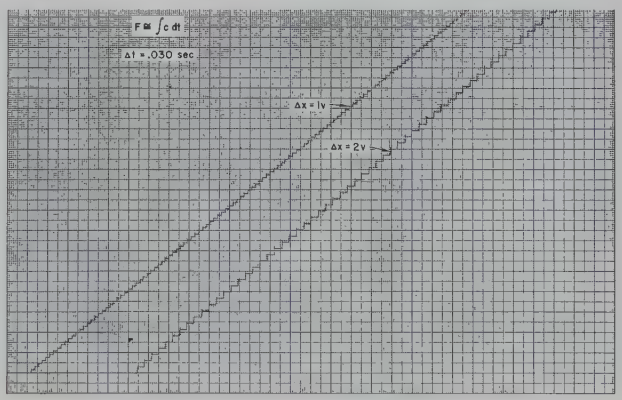


Fig. 11—Integrator output curve for Run No. 1.

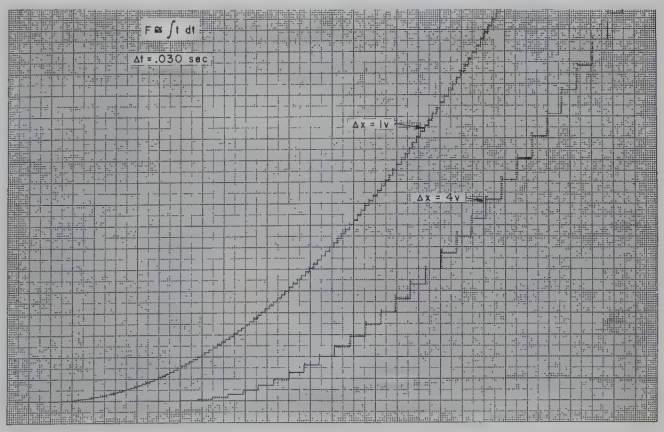


Fig. 12—Integrator output curve for Run No. 2.

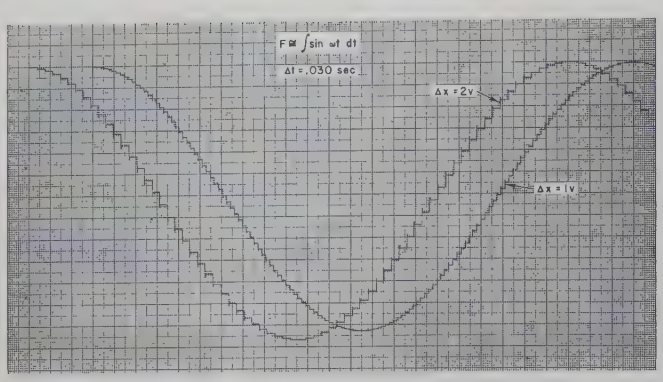


Fig. 13(a)—Integrator output curve for Run No. 3.

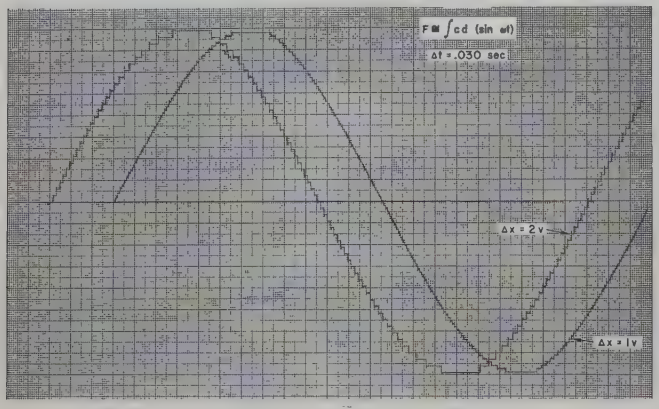


Fig. 13(b)—Integrator output curve for Run No. 4.

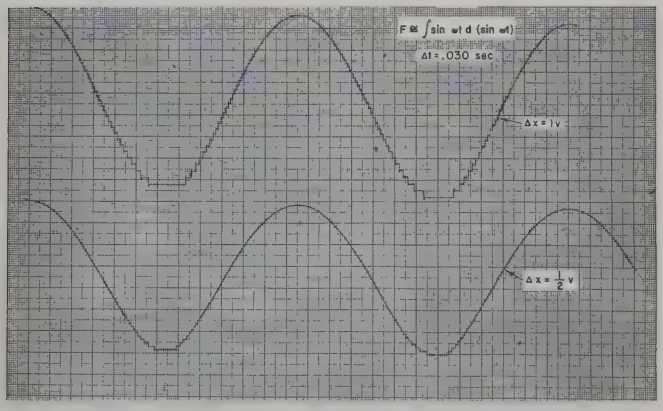


Fig. 14—Integrator output curve for Run No. 5.

Applications of a Generalized Integrator

Generation of Analytic Functions of Dependent Variables

The generation of analytic function of dependent variables, such as $\sin^2 x$, e^x, $\log x$ or $x^{2.3}$ require integration with respect to the variable x. Figs. 15(a) and 15(b) illustrate the generation of the functions x^2 and e^x. The generation of sines and cosines of dependent variables can be done without resolvers[7] by using the circuit of

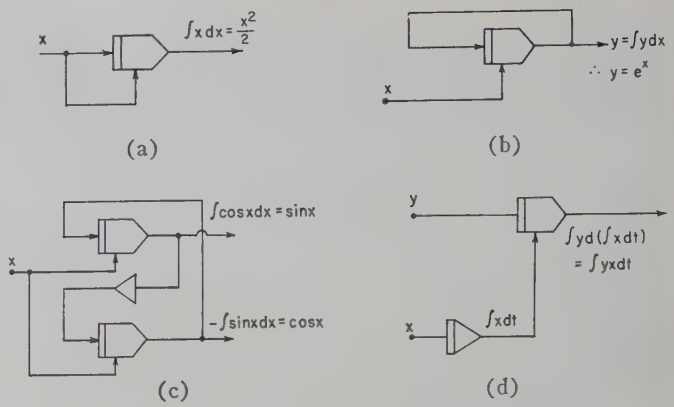


Fig. 15—Generation of analytic functions with a generalized integrator.

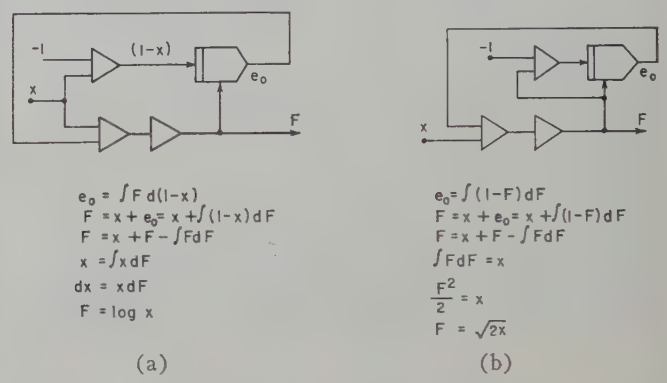


Fig. 16—Generation of analytic functions using a generalized integrator in a regenerative loop.

Fig. 15(c). A product of two variables integrated with respect to time can be obtained without a multiplier, as shown in Fig. 15(d). Similar circuits can be devised for a large number of functions which occur frequently in analog computation.

Regenerative Connections

This principle was originated by Amble[8] in connection with a mechanical differential analyzer. A single generalized integrator and two summers are required to obtain the logarithm or the square root of a variable, as indicated in Fig. 16.

Solution of Problems in Cylindrical or Spherical Coordinates

Problems expressed in such coordinate systems may require integration with respect to r, θ, and z for solution. A typical example was the study of the water-bell phenomenon,[9] where integration with respect to arc length s and altitude z were required.

[7] R. M. Howe and E. G. Gilbert, "Trigonometric resolution in analog computers by means of multiplier elements," IRE Trans. on Electronic Computers, vol. EC-6, pp. 86–91; June, 1957.

[8] O. Amble, "On a principle of connection for bush integrators," *J. Sci. Instr.* (*London*), vol. 27, pp. 284–290; December, 1956.

[9] G. N. Lance and E. C. DeLand, "On the differential analyzer solution of the water bells problem," *Proc. Roy. Phys. Soc.* (B), vol. 68, pp. 54–55; 1955.

Problem Rearrangement

Many examples could be cited from situations where rearrangement of the problem equation for simplest solution may lead to a need for integration with respect to dependent variables. This situation will almost certainly occur if a nonlinear equation is integrated once. Consider the Van der Pol equation

$$\ddot{y} + \mu(y^2 - 1)\dot{y} + y = 0. \tag{31}$$

If we integrate the equation once we obtain

$$\dot{y} + \mu \int (y^2 - 1)dy + \int y dt = C. \tag{32}$$

Direct mechanization of (31) requires a multiplier, while (32) requires a generalized integrator.

Conclusions

Integration with respect to dependent variables ordinarily requires differentiation and multiplication. A different approach to generalized integration as presented in this paper lends itself to construction with relatively few components and could probably be manufactured for a cost comparable to that of electronic multipliers. Availability of such integrators would open up new methods of problem solution which could greatly increase the versatility and applications of the analog computer.

Appendix

Differentiation with a Time Delay

Consider the differentiation of the function

$$x = A \cos \omega t \tag{33}$$

with a perfect time-delay unit. Then

$$\frac{dx}{dt} = \frac{-A \cos \omega t - A \cos \omega(t - \tau)}{\tau} \tag{34}$$

and the error can be expressed directly as

$$\epsilon = -\omega A \sin \omega t - \frac{A \cos \omega t - A \cos \omega(t - \tau)}{\tau}. \tag{35}$$

Expanding the second term on the right we obtain

$$\epsilon = A\left[-\omega \sin \omega t - \frac{\cos \omega t - (\cos \omega t \cos \omega\tau + \sin \omega t \sin \omega\tau)}{\tau}\right]. \tag{36}$$

Collecting terms we have

$$\epsilon = A\left[\sin \omega t\left(\frac{\sin \omega\tau}{\tau} - \omega\right) + \cos \omega t\left(\frac{\cos \omega\tau - 1}{\tau}\right)\right]. \tag{37}$$

Let $\omega\tau = 0.02$. Then $\sin(\omega\tau) \cong \omega\tau$, and $\cos(\omega\tau) = 0.9998$, so that

$$e \cong -\frac{.0002}{\tau} A \cos \omega t. \tag{38}$$

The error obtained in differentiating a cosine wave is thus in phase with the input wave. If the delay time $\tau = 0.02$ second, the error equals approximately 1 per cent of the input amplitude of a 1 radian per second sinewave.

Acknowledgment

The author wishes to express his appreciation to J. W. Ward of The RAND Corp., for helpful ideas and profitable discussions, and to W. Whittier of the Beckman Computation Center, for his help in running time-delay differentiation data and further profitable discussions

A Perturbation Technique for Analog Computers*

L. BUSH† AND P. ORLANDO†

Summary—A study of the motion of a fin-stabilized rocket was undertaken to determine the effect of perturbing forces on the trajectory. The mechanization of a complete problem for an analog computer to include small disturbing forces would result in trajectories which are essentially indistinguishable from the "nominal" or "unperturbed" case because of analog computer accuracy limitations. Instead, the equations of motion for the "nominal" case and the "perturbed" case, derived by first order ballistic perturbation theory, were solved simultaneously with the nominal solution providing inputs to the perturbed solution. The analog computer solution provided both the nominal trajectory and perturbations from this trajectory.

To illustrate the method, the technique is applied to the two-dimensional motion of a rocket in the vertical plane and includes perturbations due to uncertainties in winds, atmospheric density, thrust malalignments, and stability margin.

I. Introduction

ONE of the fundamental limitations in the use of analog computers is their inherent low accuracy when compared with digital techniques. Despite advances in component development, over-all accuracies better than 1 per cent are difficult to obtain. This is especially important in problems where the effects of small changes in parameters are to be observed. Fortunately there are mathematical techniques which can be applied to such cases which overcome some of these difficulties. Perhaps the most famous of these is perturbation theory, long used in celestial mechanics by Poincaré[1] and others, and more recently in the ballistics of shells and rockets by Moulton[2] and Bliss.[3] Applications to guidance problems have been made by Tsien,[4] Drenick[5] and Rosenberg.[6]

The standard technique in applying perturbation theory is to replace each variable by the sum of two variables, a nominal term and a quantity which represents a small variation from the nominal. The resulting equations are then expanded in powers of the variational quantities, and only first-order terms are retained. The final result is a set of perturbation equations whose coefficients are time-varying parameters obtained from solutions of the nominal equations.

* Manuscript received by the PGEC, November 14, 1958. This paper was presented at the National Simulation Conference, Dallas, Tex., October 23–25, 1958.

† Cornell Aeronaut. Lab., Inc., Buffalo, N. Y.

[1] H. Poincaré, "Mecanique Celeste," Paris, France, 1892.

[2] F. R. Moulton, "New Methods in Exterior Ballistics," Chicago University Press, Chicago, Ill.; 1926.

[3] G. A. Bliss, "Mathematics for Exterior Ballistics," John Wiley and Sons, Inc., New York, N. Y.; 1944.

[4] H. S. Tsien, "Engineering Cybernetics," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 13; 1954.

[5] R. Drenick, "The perturbation calculus in missile ballistics," *J. Franklin Inst.*, vol. 251, pp. 423–436; April, 1951.

[6] R. M. Rosenberg, "On flight trajectories in the neighborhood of a known trajectory," *J. Franklin Inst.*, vol. 266, pp. 109–128; August, 1958.

Once the perturbation equations have been formulated, they are solved by

1) calculating the solution to the nominal equations, and
2) applying these values step by step in the perturbation equations.

This is usually done by numerical integration of the equations and involves considerable computation, either by hand or machine. However, with the use of an analog computer it becomes feasible to solve both the nominal equations and the perturbation equations simultaneously and use the output of the nominal equations as an input to the perturbation equations. This is shown in block diagram form in Fig. 1 for the equations of motion of a ballistic rocket. Before discussing this problem in detail, the use of perturbation theory will be illustrated by a simple application.

II. A Simple Application of Perturbation Theory

Consider the motion of a mass particle in a viscous medium in the absence of gravity. Fig. 2 shows the force and kinematic diagram. The equation of motion is

$$\frac{dv}{dt} = -\frac{k}{m}v \tag{1}$$

where v is the velocity at any time t, m is the mass, and k is the retarding force per unit velocity. If v_0 is the initial velocity, the solution of (1) is easily obtained as

$$v = v_0 e^{-(k/m)t}. \tag{2}$$

If we now wish to determine the perturbation in velocity due to a slight change in particle mass, say 0.1 per cent, it becomes impractical to solve (1) with the new mass. Instead, a perturbed equation may be obtained as follows. In (1) v is replaced by $\bar{v}+\delta v$ and m by $\bar{m}+\delta m$, where $\bar{v}$ and $\bar{m}$ are the nominal values and δv, δm are perturbations from the nominal. The resulting equation is expanded in powers of small quantities and only the first-order terms are retained.

$$\frac{dv}{dt} = F(v, m)$$

$$\frac{d\bar{v}}{dt} + \delta\left(\frac{dv}{dt}\right) = F(\bar{v}, \bar{m}) + \frac{\partial F}{\partial v}\delta v + \frac{\partial F}{\partial m}\delta m \tag{3}$$

where

$$\frac{\partial F}{\partial v} = -\frac{k}{\bar{m}}; \qquad \frac{\partial F}{\partial m} = \frac{k\bar{v}}{\bar{m}^2}\, d\,.$$

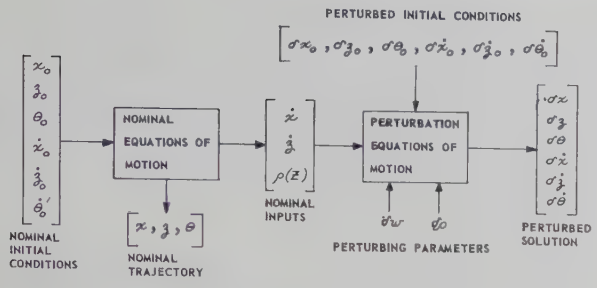


Fig. 1—Technique of mechanization.

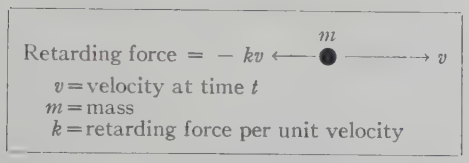


Fig. 2—Particle in viscous medium.

Eq. (1) is now subtracted from (3) to obtain the perturbation equation,

$$\frac{d}{dt}(\delta v) + \frac{k}{\bar{m}}(\delta v) = \frac{k\delta m}{\bar{m}^2}\bar{v}\,. \tag{4}$$

The solution is

$$\delta v = \bar{v}k\frac{\delta m}{\bar{m}^2}t. \tag{5}$$

The perturbed solution differs from the nominal solution by the factor

$$\left(k\frac{\delta m}{\bar{m}^2}t\right).$$

Fig. 3 shows both the nominal and perturbed solutions as dimensionless plots. Note that the maximum perturbation is less than 0.04 per cent of the initial velocity. Imagine trying to notice this effect on a conventional analog computer simulation. The perturbation simulation technique could be applied to this particular problem by simulating both (1) and (4), using the nominal velocity output of(1) $\bar{v}$, as an input to (4).

III. Application to Rocket Dispersions

A. *Nominal and Perturbation Equations of Motion*

Only motion in the vertical plane for a nonrolling rocket after burnout is considered here. Similar techniques may be applied to a study of the equations of motion before burnout. The force and kinematic diagram is shown in Fig. 4 where the following definitions are used:

x, z = earth fixed axis system with x horizontal, z vertical and origin at the launcher.
θ = rocket attitude angle above horizontal.
α = angle of attack.
V = airspeed.

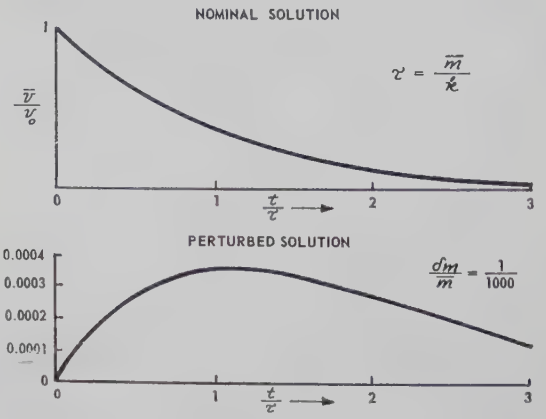


Fig. 3—Nominal and perturbed solutions.

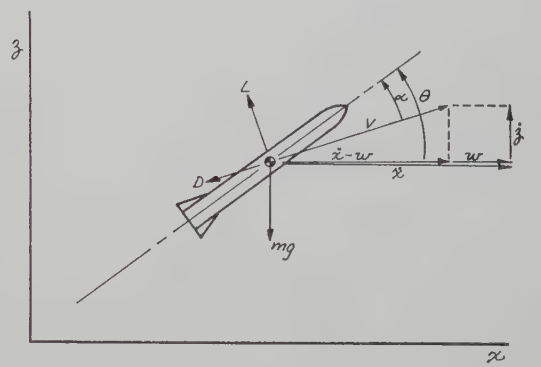


Fig. 4—Force and kinematic diagram for rocket after thrust cutoff.

w = horizontal wind velocity, positive in the x-direction.
$\dot{x}$, $\dot{z}$ = components of velocity with respect to ground.
m = rocket mass.
g = acceleration due to gravity.
L = lift force normal to airspeed V.
D = drag force directed opposite to V.

The nominal equations of motion for the rocket are (see Appendix):

$$\ddot{x} = -\frac{1}{2}\frac{\rho VS}{m}\left[\dot{z}C_L + (\dot{x} - w)C_D\right]$$

$$\ddot{z} = -\frac{1}{2}\frac{\rho VS}{m}\left[\dot{z}C_D - (\dot{x} - w)C_L\right] - g$$

$$\ddot{\theta} = \frac{1}{2}\frac{\rho V^2 Sb}{I}C_M$$

where

$$V = \left[\dot{z}^2 + (\dot{x} - w)^2\right]^{1/2}, \text{ the airspeed}$$

and

$$\alpha = \theta - \tan^{-1}\frac{\dot{z}}{\dot{x} - w}, \text{ the angle of attack.}$$

The additional symbols appearing in these equations are:

ρ = atmospheric density.
S = reference area.
I = lateral moment of inertia.
b = reference length.
C_L = lift coefficient.
C_D = drag coefficient.
C_M = moment coefficient.

Using the techniques described previously, the perturbation equations are then derived for the case where effects of slight changes in wind (δw) and density ($\delta\rho$) are to be investigated (see Appendix). These are:

$$\delta\ddot{x} = -\frac{1}{2}\frac{\rho VS}{m}\left\{\frac{\dot{x}}{\rho}C_{D_0}\delta\rho + \left(\frac{\dot{z}^2}{V^2}C_{L_{\alpha_0}} + \frac{V^2+\dot{x}^2}{V^2}C_{D_0}\right)\delta\dot{x} + \frac{\dot{x}\dot{z}}{V^2}(C_{D_0} - C_{L_{\alpha_0}})\delta\dot{z} + \dot{z}C_{L_{\alpha_0}}\delta\theta - \left(\frac{\dot{z}^2}{V^2}C_{L_{\alpha_0}} + \frac{V^2+\dot{x}^2}{V^2}C_{D_0}\right)\delta w\right\}$$

$$\delta\ddot{z} = -\frac{1}{2}\frac{\rho VS}{m}\left\{\frac{\dot{z}}{\rho}C_{D_0}\delta\rho - \frac{\dot{x}\dot{z}}{V^2}(C_{L_{\alpha_0}} - C_{D_0})\delta\dot{x} + \left(\frac{\dot{x}^2}{V^2}C_{L_{\alpha_0}} + \frac{V^2+\dot{z}^2}{V^2}C_{D_0}\right)\delta\dot{z} - \dot{x}C_{L_{\alpha_0}}\delta\theta + \frac{\dot{x}\dot{z}}{V^2}(C_{L_{\alpha_0}} - C_{D_0})\delta w\right\}$$

$$\delta\ddot{\theta} = \frac{1}{2}\frac{\rho Sb}{I}C_{M_{\alpha_0}}\{\dot{z}\delta\dot{x} - \dot{x}\delta\dot{z} + V^2\delta\theta - \dot{z}\delta w\}.$$

The symbols are defined as for the nominal equations. In addition we have $\delta(\ \ )$, perturbation of a parameter from its nominal value,

$$C_{L_{\alpha_0}} = \left(\frac{\partial C_L}{\partial\alpha}\right)_{\alpha=0}$$

$$C_{M_{\alpha_0}} = \left(\frac{\partial C_M}{\partial\alpha}\right)_{\alpha=0}$$

$$C_{D_0} = (C_D)_{\alpha=0}.$$

Note that time-varying parameters of the solution to the nominal equations (ρ, V, $\dot{x}$, $\dot{z}$,) appear as coefficients in the perturbation equations. For the nominal trajectory the angle of attack is essentially zero after thrust cutoff and the aerodynamic coefficients are evaluated for this condition.

B. *Analog Computer Mechanization*

The initial conditions for the nominal equations starting at thrust cutoff were known from previous calculations, as well as the initial conditions for the perturbation equations due to thrust malignment. It was assumed that other perturbations during the boost phase could be overcome by a rudimentary guidance system. Thus only the effects of perturbing forces following thrust cutoff were considered. Fig. 1 indicates the technique used for mechanization. The nominal velocity components and density-altitude function served to determine the variable coefficients of the perturbation equations. Solutions to the perturbation equations were obtained for the following perturbing inputs applied independently:

1) Range wind
2) Density variation from standard atmosphere
3) Perturbed initial conditions due to thrust malalignment
4) Variation in stability margin. This was done by changing C_M.

Fig. 5 shows the effect of range wind on the range perturbation for various types of trajectories. Three trajectories are included, the maximum range and the one half-maximum range for both high and low angle launchings. The range perturbation is expressed in mils, or parts per thousand parts of nominal range.

Fig. 6 shows the effect of percentage variation in atmospheric density on the range for the same three types of trajectories. Fig. 7 shows the effect of thrust malalignment for the maximum range trajectory only. Changes in C_M of ± 50 per cent resulted in insignificant perturbations in range. Note that in all of these curves, the quantities being investigated involve extremely small values.

IV. Conclusion

It has been shown that conventional perturbation techniques can be used simultaneously with the solution of a system of nominal equations on an analog computer to obtain improved accuracy. Although the utility of the method has been illustrated by its application to rocket dispersion problems, the real significance lies in its generality. This method of approach can have wide applications to problems where small changes in parameters affect the nominal solution, without resorting to extensive digital computation.

Appendix

Derivation of Rocket Nominal and Perturbation Equations

Referring to Fig. 4, the following equations may be written:

a) Kinematic relations:

$$\alpha = \theta - \tan^{-1}\frac{\dot{z}}{\dot{x}-w} \qquad \sin(\theta-\alpha) = \frac{\dot{z}}{V}$$

$$V = [\dot{z}^2 + (\dot{x}-w)^2]^{1/2} \qquad \cos(\theta-\alpha) = \frac{\dot{x}-w}{V}.$$

b) Dynamical Equations:

$$m\ddot{x} = -L\sin(\theta-\alpha) - D\cos(\theta-\alpha)$$

$$m\ddot{z} = L\cos(\theta-\alpha) - D\sin(\theta-\alpha) - mg$$

$$I\ddot{\theta} = M$$

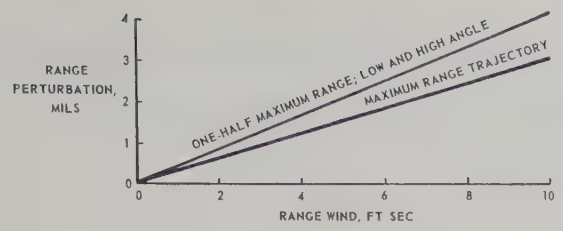


Fig. 5—Effect of range wind.

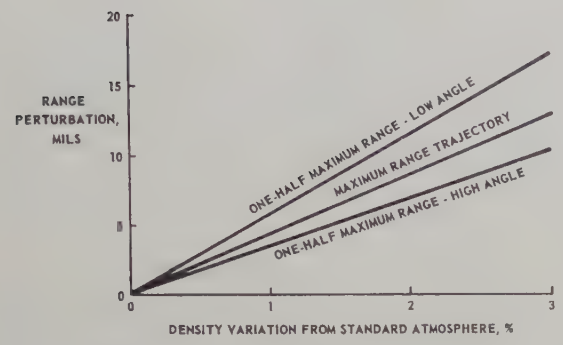


Fig. 6—Effect of density variation.

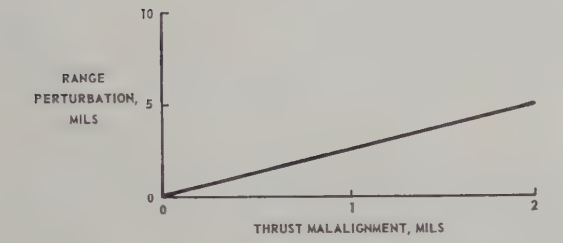


Fig. 7—Effect of thrust malalignment for maximum range trajectory.

where L is the lift force, D is the drag force, and M is the total moment acting on the rocket. The other symbols are as previously defined. After introducing aerodynamic coefficients,

$$C_L = \frac{L}{\frac{1}{2}\rho V^2 S}, \quad C_D = \frac{D}{\frac{1}{2}\rho V^2 S}, \quad \text{and} \quad C_M = \frac{M}{\frac{1}{2}\rho V^2 S b},$$

the complete system of equations appears as given in Section III. We assume that C_L, C_D, and C_M are functions of α only, ρ is a function of z only, and I, b, S, M, and g are constants. To obtain the perturbation equations, the procedure outlined in Tsien[4] is used. The nominal solution is affected by perturbations in ρ, $\dot{x}$, $\dot{z}$, θ, and w. The dynamical equations are rewritten as

$$\ddot{x} = F(\rho, \dot{x}, \dot{z}, w, C_L, C_D),$$
$$\ddot{z} = G(\rho, \dot{x}, \dot{z}, w, C_L, C_D), \text{ and}$$
$$\ddot{\theta} = H(\rho, \dot{x}, \dot{z}, C_M);$$

and the perturbation equations are, to first-order accuracy,

$$\delta\ddot{x} = \frac{\partial F}{\partial \rho}\delta\rho + \frac{\partial F}{\partial \dot{x}}\delta\dot{x} + \frac{\partial F}{\partial \dot{z}}\delta\dot{z} + \frac{\partial F}{\partial w}\delta w,$$

with similar expressions for $\delta\ddot{z}$, and $\delta\ddot{\theta}$. In taking the partial derivatives of the indicated functions it must be remembered that V and α are functions of $\dot{x}$, $\dot{z}$, w, and θ, as indicated by the kinematic relations. The partial derivatives are to be evaluated for the nominal or unperturbed path. Inserting the partial derivatives enables one to arrive at the perturbation equations presented in Section III.

A Four-Quadrant Multiplier Using Triangular Waves, Diodes, Resistors, and Operational Amplifiers*

PAUL E. PFEIFFER†

Summary—**A simple scheme of switching triangular waves and measuring the average current through resistors into a low impedance summing point makes possible four-quadrant multiplication with four diodes, precisely adjusted resistors, and a means of measuring the current. A practical circuit utilizes one operational amplifier to obtain $-(X+Y)/2$ and a second such unit to measure the summing point current. Addition of four auxiliary diodes reduces circuit interactions and makes less stringent requirements on the diodes. A simple operational adjustment procedure is described. Also, a simple means for obtaining the precise resistance balance is outlined. Calibration does not depend upon triangular wave frequency or symmetry. The amplifiers are not required to handle the triangular wave frequencies.**

THE operation of the multiplier described in this paper is based upon a simple squaring system utilizing triangular waves and upon a modification of the quarter-squares method of multiplication. The method of squaring is based on the mathematical fact that the areas of similar triangles are proportional to the squares of their heights. In the first four sections, the mathematical basis of the method will be developed and circuit behavior will be described on the assumption that the elements perform in an ideal manner. In subsequent sections, practical considerations arising from nonideal behavior of the elements will be discussed and certain experimental results will be noted.

A Mathematical Squaring System

The manner in which the basic property of triangular waves is utilized may be understood with the aid of the diagram of Fig. 1. In that figure there is shown a triangular wave having peak to peak swing of $2A$. The wave has been shifted an amount s from the position of equal positive and negative swings. The portion of the shifted wave above the zero reference will be referred to as the upper wave and the portion below the reference will be called the lower wave. Suppose the letter U represents the average ordinate of the upper wave, the letter L represents the average ordinate of the lower wave, and the letter S represents the ratio s/A. Then, simple geometrical considerations show that

$$U = \frac{A}{4}(1 + s/A)^2 = \frac{A}{4}(1 + S)^2 \text{ for } |S| \leq 1 \quad (1)$$

and

$$L = -\frac{A}{4}(1 - s/A)^2 = -\frac{A}{4}(1 - S)^2 \text{ for } |S| \leq 1. \quad (2)$$

* Manuscript received by the PGEC, November 3, 1958; revised manuscript received, February, 16, 1959. This paper was presented at the National Simulation Conference, Dallas, Tex., October 23–25, 1958.

† Dept. of Elect. Engrg., Rice Institute, Houston, Tex.

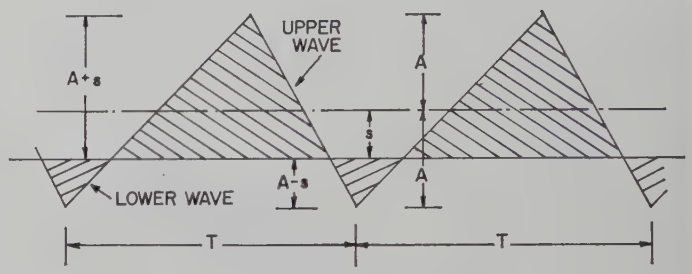


Fig. 1—Triangular wave or Δ-wave.

Fig. 1 is drawn for s positive, but the expressions hold for s or S of either sign. It should be noted that the average ordinate does not depend upon the period T of the triangular wave nor on the relative slopes of the leading and trailing edges.

Diode Switching Circuits to Perform the Squaring Operation

Shifting of the triangular wave and determination of the average ordinates for the upper and lower waves may be accomplished electrically by the simple circuits shown in Figs. 2(a) and 2(b), respectively. The symbol Δ refers to the triangular wave, presented physically as a voltage wave. The letters x and y refer to voltage inputs which provide the shift of the triangular wave. Let e_Δ, e_x, and e_y designate the instantaneous voltages corresponding to the triangular wave Δ and the quantities x and y, respectively. Let e_d be the voltage at the junction point of the three resistors in each of the two circuits. This voltage is the voltage across the diodes. If ideal diodes are assumed, there are two conditions of operation:

1) Diode nonconducting: $i = 0$.
2) Diode conducting: $e_d = 0$.

For the nonconducting condition, the voltage relations are

$$\frac{e_\Delta - e_d}{R} + \frac{e_x - e_d}{R/a} + \frac{e_y - e_d}{R/b} = 0 \quad (3)$$

Simple algebra suffices to reduce this to the form

$$e_d(1 + a + b) = e_\Delta + ae_x + be_y. \quad (4)$$

For the condition in which the diodes are conducting

$$i = (e_\Delta + ae_x + be_y)/R. \quad (5)$$

For the circuit configuration of Fig. 2(a), the current i_1 flows whenever the diode voltage e_d is positive, which is the same as the condition that the right hand side of

(5) is positive. If the peak to peak swing of the triangular voltage wave is $2E$, the i_1 current wave is just the upper wave of Fig. 1, with

$$A = E/R \text{ and } s = (ae_x + be_y)/R = \frac{E}{R}(aX + bY) \quad (6)$$

where $X = e_x/E$ and $Y = e_y/E$.

From the geometry of Fig. 1, it follows that

$$\text{Average } i_1 = \frac{E}{4R}(1 + aX + bY)^2 \quad (7)$$

when $|aX+bY| \leq 1$.

The peak current attainable is E/R.

In an exactly similar manner, it can be argued that for the circuit of Fig. 2(b) the average current is given by

$$\text{Average } i_2 = -\frac{E}{4R}(1 - aX - bY)^2 \quad (8)$$

when $|aX+bY| \leq 1$.

It should be apparent that the circuits can be reduced to the case of a single variable X or Y by reducing one of these quantities to zero or by reducing the proper multiplier to zero. The latter operation is achieved by making the resistance infinite by opening a switch or removing the resistor.

A Modified Quarter-Squares Combination for Multiplication

Several different combinations of squared terms of the type appearing in (7) and (8) lead to the product of X and Y. The combination leading to the simplest circuit configurations discovered so far is the following:

$$(1 + X/2)^2 + (1 + Y/2)^2 - [1 - (-X - Y)/2]^2 - 1 = -XY/2. \quad (9)$$

The circuit of Fig. 3 will achieve this combination, provided the device to measure average current is of sufficiently low impedance that it does not produce appreciable voltage interaction at the summing point. For this circuit,

$$\text{Average } i = -\frac{E}{8R}XY = -\frac{e_x e_y}{8ER} \quad (10)$$

for $|X| \leq 1$, $|Y| \leq 1$.

A Circuit Using Operational Amplifiers

The circuit of Fig. 3 requires that both x and $-x$ as well as both y and $-y$ be available. A means of measuring the average current must be provided. All of the input voltages must be provided from low impedance sources so that the precise resistance balance is not disturbed. A modification of the circuit of Fig. 3 is shown in Fig. 4. Two operational amplifiers are used. The first amplifier provides $-(X+Y)/2$. The second amplifier provides an output voltage proportional to the summing point current. The network associated with the feedback resistor serves to filter out the triangular wave frequencies, so that the output voltage corresponds to the average current. This means that the amplifier does not have to respond to the high-frequency components of voltages produced by the rectified triangular waves. If the voltages e_x and e_y vary slowly with respect to the variations due to the triangular waves and the rapid variations due to the switched triangular waves are filtered out, the output voltage of the second amplifier becomes

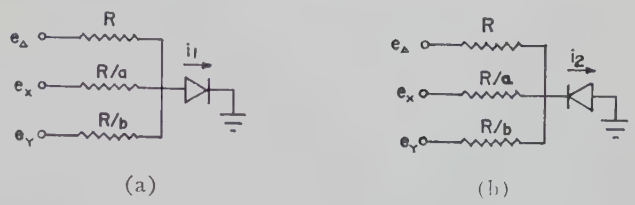


Fig. 2—Diode-switching circuits to perform the squaring operation; (a) upper wave current, (b) lower wave current.

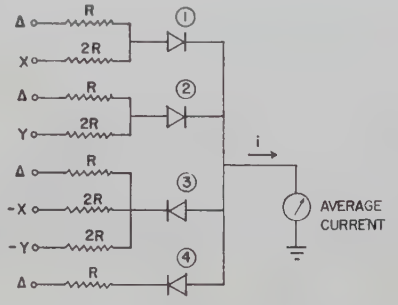


Fig. 3—Combination of diode-switching circuits which perform multiplication.

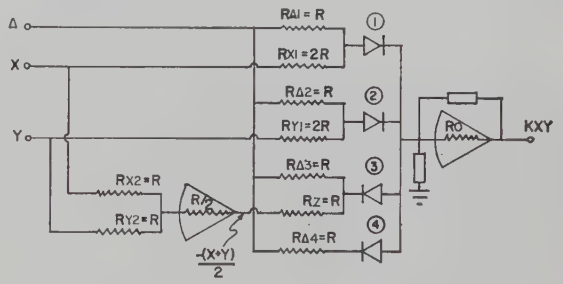


Fig. 4—A modification of the circuit of Fig. 2, utilizing operational amplifiers.

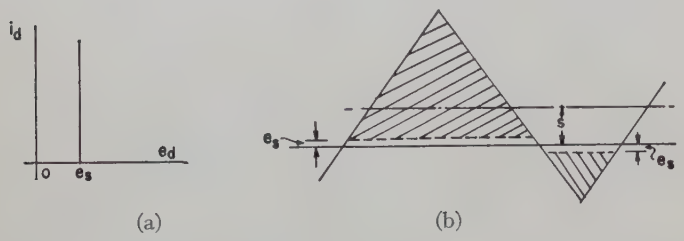
$$e_0 = (ER_0/8R)XY = (R_0/8ER)e_x e_y. \quad (11)$$

Proper adjustment of the ratio R_0/R and suitable choice of the peak triangular voltage E serve to give the desired scale factor.

The previous description of the circuit has assumed perfect elements, a perfect triangular wave, and steady values of the voltages e_x and e_y. If the triangular carrier wave frequency is sufficiently high that the voltages e_x and e_y are essentially constant throughout any given cycle, the multiplier is suitable for dynamic multiplication. In the following discussion of the errors due to various causes, it is assumed that carrier frequency is sufficiently high that dynamic errors are negligible.

The Effect of Nonzero Switching Voltage for the Diodes

One of the serious practical problems in the design and adjustment of the multiplier system arises from the fact that the diodes do not switch at zero voltage. The exact shape of the current versus voltage curve can be determined experimentally. Experiment has shown, however, that for the switching operation, the effect of nonzero switching may be analyzed to a first approximation with the aid of the idealized model of Fig. 5.



(a) (b)

Fig. 5—Nonzero switching voltage for diodes and the effects on upper and lower waves of current into the summing point; (a) assumed diode characteristics, (b) effect of e_s on upper and lower waves.

The diodes are considered ideal except for an offset of the switching voltage e_s from the zero value. The effect of this offset is to alter the switching level. If $\epsilon = e_s/E$, then (7) and (8) are modified to give, respectively,

$$\text{Average } i_1 = \frac{E}{4R}(1 + aX + bY - \epsilon)^2 \qquad (12)$$

and

$$\text{Average } i_2 = -\frac{E}{4R}(1 - aX - bY - \epsilon)^2. \qquad (13)$$

The quarter-squares combination represented by the circuits of Fig. 3 and Fig. 4 becomes upon expansion

$$\begin{aligned}\text{Average } i = -\frac{E}{8R}\,[&XY + 4(\epsilon_1 + \epsilon_2 - \epsilon_3 - \epsilon_4) \\ &+ 2(\epsilon_1 - \epsilon_3)X + 2(\epsilon_2 - \epsilon_3)Y \\ &- 2(\epsilon_1^2 + \epsilon_2^2 - \epsilon_3^2 - \epsilon_4^2)]. \end{aligned} \qquad (14)$$

The subscripts on the epsilons correspond to the numbers of the diodes. The higher-order error terms ordinarily can be neglected. Probably the most important point to be noted in this expression is that if all of the diodes are balanced with respect to effective switching voltages, the error due to the switching voltage is reduced to zero. This fact is important in the case of crystal diodes, since the value of e_s may be of the order of one-quarter volt. The symmetry of the circuit leads to a balancing out of these voltages, provided they are all the same.

Errors Due to Resistance Unbalance

It is convenient to divide the resistance unbalances into two kinds. One type of unbalance is that between groups associated with the various diodes. These may be accounted for by considering the way the ratios of the resistors to which the triangular voltage is applied differ from unity. If these resistors are denoted by $R_{\Delta k}$ where k is the number of the diode in Fig. 3 or Fig. 4, then it is convenient to put

$$R_{\Delta 3} = R \quad \text{and} \quad R_{\Delta k} = R/(1 + b_k) \qquad (15)$$

for $k = 1$, 2, and 4.

The second type of unbalance is due to inexact ratios within the group tied to each diode. This can be accounted for as a variation of the factors multiplying X, Y, and $(X+Y)$ in the terms in the expressions for average current. Since ideally these factors are all one-half, the unbalance may be accounted for by writing the factors as $(1+c_k)/2$. The quantities b_k and c_k are measures of the amount of unbalance. The expression for average current, under the assumption of perfect diode switching, becomes

$$\begin{aligned}\text{Average } i = \frac{E}{4R}\,\{&(1 + b_1)[1 + (1 + c_1)X/2]^2 \\ &+ (1 + b_2)[1 + (1 + c_2)Y/2]^2 \\ &- [1 + (1 + c_3)(X + Y)/2]^2 - (1 + b_4)\}. \end{aligned} \qquad (16)$$

Upon expansion, this becomes, after neglecting some second-order terms,

$$\begin{aligned}\text{Average } i \doteq -\frac{E}{8R}\,\{&XY(1 + 2c_3) - 2(b_1 + b_2 - b_4) \\ &- 2(b_1 + c_1 - c_3)X - 2(b_2 + c_2 - c_3)Y \\ &- (b_1 + 2c_1 - 2c_3)X^2/2 \\ &- (b_2 + 2c_2 - 2c_3)Y^2/2\}. \end{aligned} \qquad (17)$$

Eq. (17) exhibits the effects of various unbalance parameters. It is of interest to ask whether an appropriate condition of unbalance could be obtained to counteract the error terms due to switching voltage as presented in (14). For one thing, the constant term could be balanced by a suitable adjustment of b_4 by modifying $R_{\Delta 4}$ appropriately. The fact that the other unbalance parameters appear in coefficients of X^2 or of Y^2 or both indicates that it is not wise to try to balance out the terms in (14) which are linear in X or Y. While this balance could be made accurately at certain points, there would be considerable tracking error. Experience has shown that the resistors may be balanced closely enough to keep the unbalance factors of the order of 0.01 per cent. This makes it seem wise to balance all the resistors, except possibly $R_{\Delta 4}$, on a factory adjustment basis or at least on a maintenance basis and to seek to provide operational adjustments in some other manner. Because of the good resistance balance obtainable, further analysis will be carried out on the assumption that all unbalance paramenters except b_4 are negligibly small and may thus be considered to be zero.

A Method of Balancing the Multiplier Circuit

A combination of analysis and experiment has led to the following method of balancing the multiplier offset. The X-zero product shift and the Y-zero product shift are balanced by means of constant terms added to the first and second quantities in parenthesis, respectively, in (9). Suppose an offset δ_1 is added to $X/2$ and an offset δ_2 is added to $Y/2$ in the terms referred to. This may be done for $X/2$ by applying a voltage through a third resistor to diode number 1 in Fig. 3, or its modifications. The offset for $Y/2$ may be achieved similarly by a voltage and resistor connected to diode number 2. The resulting equation for average current becomes

$$\begin{aligned}\text{Average } i = -\frac{E}{8R}\{&XY + 4[\epsilon_1 + \epsilon_2 - \delta_1 - \delta_2 \\ &- \epsilon_3 - \epsilon_4 + b_4(1-\epsilon_4)^2/2] \\ &+ 2[(\epsilon_1-\delta_1)^2 + (\epsilon_2-\delta_2)^2 - \epsilon_3^2 - \epsilon_4^2 \\ &+ 2(\epsilon_1-\delta_1-\epsilon_3)X + 2(\epsilon_2-\delta_2-\epsilon_3)Y\}. \quad (18)\end{aligned}$$

This equation shows that the X-zero product shift may be cancelled by adjusting δ_1 and the Y-zero product shift may be cancelled by adjusting δ_2. The resulting zero offset when both inputs are zero may then be cancelled by adjusting $R_{\Delta 4}$ (and hence b_4). Experiment indicates that these adjustments can be made to better than 0.1 per cent of full scale if the triangular wave is good enough. That is, the output is less than 0.1 per cent of full scale when either or both of the inputs is at zero. The nature of the multiplier operation is such that when the zero offset is cancelled, static accuracy of multiplication is assured. Slight adjustments in scale factor can be made by adjusting the triangular wave amplitude E. It should be noted that the effect on scale factor is inversely related to amplitude E, as is shown in the last member of (10). The form for the second member is misleading, since it tends to conceal the fact that X and Y are related to E.

If diodes number 1 and 2 are sufficiently well balanced, so that ϵ_1 and ϵ_2 are essentially the same, the zero-product shift adjustment can be made with a single correction δ added to $-(X+Y)/2$. In this case, δ is adjusted for the best compromise on X-zero product shift and Y-zero product shift. The resistor $R_{\Delta 4}$ is then adjusted to balance out the zero offset.

The Use of Catcher Diodes

The circuits of Fig. 3 and Fig. 4 have several disadvantages. When a diode is conducting, the low impedance to ground prevents appreciable interaction between the circuits connected to the diode. When the diode is not conducting, however, there is considerable interaction. The maximum value of the voltage across the diode may approach the peak value of the triangular wave for certain conditions of operation. If appreciable capacitance to ground exists at the junction point, the time delay due to displacement charge may introduce serious switching errors. Also, the triangular wave source and the X and Y sources see widely varying loads. Experience has shown that for values of R of 0.1 megohm in the circuits of Fig. 3 and Fig. 4 it is necessary to keep source impedances to the order of ten ohms or less. These difficulties can be ameliorated by the simple remedy shown in the circuit of Fig. 6. Each of the

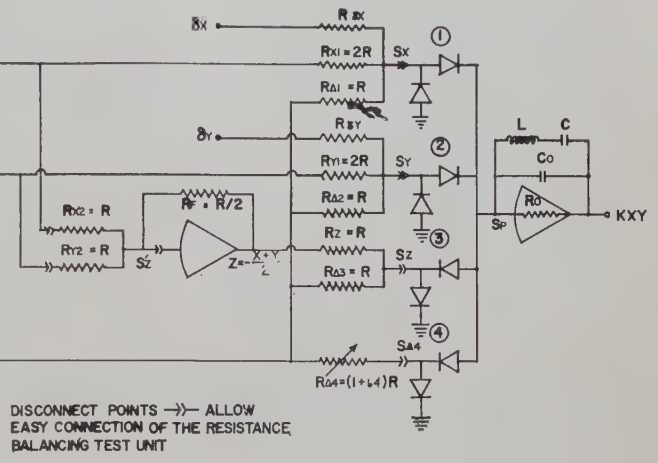


Fig. 6—Complete multiplier system with offset applied to x and to y.

principal diodes is provided with an auxiliary "catcher" diode, which prevents the back voltage on the principal diode from exceeding the switching voltage of the catcher diode. This is about one-quarter of a volt for silicon diodes. Circuit interactions are markedly reduced, diode voltage and back resistance requirements are reduced, and capacitive effects are reduced. Source resistances of the order attained in cathode follower circuits may be tolerated if these resistances do not vary appreciably over the cycle of operation. Diode requirements may be reduced to the point that eight diodes may be used which have a total cost less than the cost of four diodes which meet the more stringent requirements of the circuit of Fig. 4.

Operational Amplifier Considerations

The amplifier used to obtain $-(X+Y)/2$ is operated in a normal fashion and involves no particular problems. To prevent drift, it is desirable to use a chopper stabilized unit. The choice of the factor $1/2$ in $(X+Y)/2$ is dictated by the fact that this term must not exceed unit magnitude (in machine units) if the inputs X and Y do not.

The second amplifier, used to measure average diode current, is operated under conditions that lead to several problems. For one thing, it is necessary to keep the summing point impedance small so that there is negligible summing point voltage interaction. At low frequencies, where the amplifier gain is high, the input impedance is approximately the feedback impedance divided by the amplifier gain. At higher frequencies, the gain drops off. It is necessary to make the feedback imped-

ance drop off to prevent an untenable rise in summing point impedance.

A second requirement that must be met by the output amplifier configuration is that the triangular wave must be filtered out. This is desirable from the standpoint of the equipment tied to the output. Also, it is necessary to reduce the triangular wave content of the output voltage to prevent voltage overloading of the output amplifier. Such filtering is obtained by making the feedback impedance low at the frequency of the triangular wave and its harmonics. Fortunately, the two requirements are compatible. One solution is shown in the diagram of Fig. 6.

The value of the feedback resistor R_0 in Fig. 6 is determined with the aid of (10), which may be rewritten in terms of the output voltage e_0 as follows

$$e_0 = \frac{e_x e_y}{8E} \cdot \frac{R_0}{R} = K e_x e \,. \tag{19}$$

If the voltages are expressed in machine units, and if it is desired that $e_0 = 1$ when $e_x = e_y = 1$, then it follows that

$$\frac{R_0}{R} = 8E. \tag{20}$$

Since E must be greater than unity, it is necessary to have R_0/R greater than eight. The value of E probably should be betwen 1.1 and 1.25 machine units.

Values of R_0/R as large as that indicated by (20) may lead to undesirably high input impedances. Also, the effective feedback factor is rather small, so that amplifier drift may be a problem. Drift is usually negligible if a chopper stabilized unit is employed. It may, however, be desirable to operate with a multiplier factor K in (19) somewhat less than unity.

Results

It is apparent that the method hinges upon the availability of a precise triangular wave. The source must be stable and have relatively low output impedance—say of the order of 100 ohms or less. Although the exposition has treated the triangular wave generator as a detail, the design of a good, economical, stable unit is not a trivial matter.

Experimental work was carried out with a commercially-available triangular-wave generator whose frequency was limited to approximately 1200 cps. This low frequency made it possible to work without undue concern for capacitive effects and made it possible to isolate such effects as those due to diode switching voltages, and to resistance unbalance. It leaves open the question of high-frequency behavior. Adequate filtering of the low-frequency triangular wave introduces sizeable phase errors at signal frequencies in the range of 10–20 cps. With the use of the catcher diodes, it should be possible to increase the triangular wave source to at least 25 kc. This remains to be tried, however.

Experimental results on a prototype model have yielded zero-product shift values of the order of 0.1 per cent. Departures from linearity over four quadrants have been of the order of 0.5 per cent. It should be possible with improved mechanical design to reduce the latter figure by a factor of at least two.

A simple routine for balancing the resistors has been worked out by W. R. Peters. The auxiliary switching circuit to be built into a test unit and the balancing procedure is described in the Appendix.

Appendix

A Method for Balancing the Multiplier Resistors

The resistors in the multiplier may be balanced by a simple but effective procedure. Two resistors which are to be balanced against each other are used as the input and feedback resistors, respectively, in an ordinary inverter circuit. A fixed input test voltage E_T is applied to the input and the output is monitored by a voltmeter. The sensitivity of this voltmeter may be greatly increased by comparing the output with a "pedestal voltage" of approximately $-E_T$. Balance of the two resistors is determined by interchanging the role of the input and feedback resistors. This is done by reversing the connections of the end points of the feedback configuration—output end to input and vice versa—while leaving the junction connected to the summing point. When the resistors are balanced, the output reading should be precisely the same in the two conditions.

The circuit of Fig. 7 shows the resistance-diode portion of the multiplier circuit of Fig. 6 connected to a test circuit which provides the necessary switching to balance the resistors according to the following schedule.

Position	*Resistors to be Balanced*	*End Points*	*Junction Point*
1	$R_{\Delta 3}$ vs. $R_{\Delta 1}$	S_X, S_Z	Δ
2	$R_{\Delta 3}$ vs. $R_{\Delta 2}$	S_Y, S_Z	Δ
3	$R_{\Delta 3}+R_{\Delta 2}$ vs. R_{Y1}	Y, S_Z	S_Y
4	$R_{\Delta 3}+R_{\Delta 1}$ vs. R_{X1}	X, S_Z	S_X
5	R_{X2} vs. R_{Y2}	X_2, Y_2	S_Z'

The circuit of Fig. 7 is drawn with the selector switch in position 3, for which R_{Y1} is balanced against the series combination of $R_{\Delta 3}+R_{\Delta 2}$. The junction is the point marked S_Y. One of the poles of the rotary switch connects this junction to the amplifier summing point. One set of the end points is connected in common. The other end point for each test is selected by the other pole of the rotary switch. The end points are connected to the amplifier output and to the test voltage source through

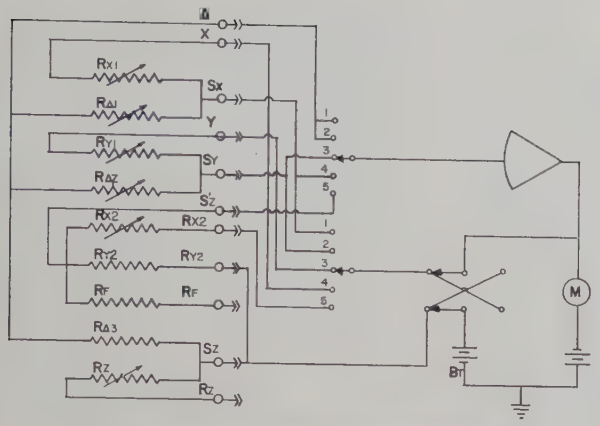


Fig. 7—Resistance balancing circuit—part one.

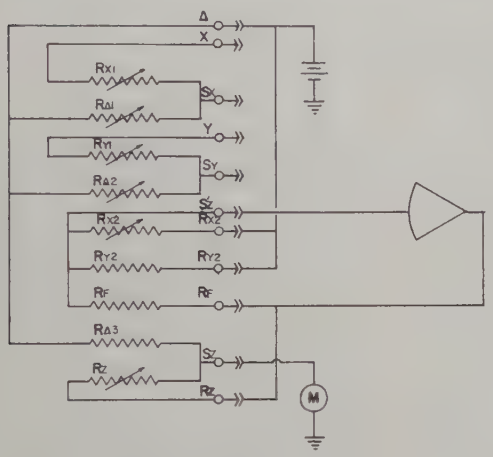


Fig. 8—Resistance balancing circuit—part two.

the reversing switch. Rapid balance to better than 0.01 per cent may be achieved by using E_T of approximately 45 volts dc.

The balancing procedure is not completed in the five steps obtained with the setup of Fig. 7. It is still necessary to make the transfer admittance from the point marked Δ to S_Z equal in magnitude and opposite in sign to twice that from either X or Y to S_Z. This is accomplished by paralleling R_{X2} and R_{Y2} so that the transfer admittance from the combined point to S_Z should have the same magnitude as that from Δ to S_Z. Points Δ, X, and Y are tied together; E_T is applied; R_Z is adjusted so that a null is observed at the junction point marked S_Z. The circuit of Fig. 8 provides this configuration. It is possible to combine the circuits of Fig. 7 and Fig. 8 by suitable switching. It is simpler to separate the two circuits by having separate plugs in the test unit. Refinements in the switching to protect the meter may also be desirable. If so, they are easily provided.

Acknowledgment

The author is indebted to a number of students who have carried out some of the experimental work on the project. In particular, Betsalel Tasini, who, while a graduate student, constructed and tested a model of the multiplier based upon a more complicated circuit than the one described. Most of the experimental work on the circuit has been carried out by W. R. Peters, electronics technician in the Department of Electrical Engineering, Rice Institute. He is due full credit for the resistance-balancing scheme described in the Appendix.

Correspondence

Dual-Polarity Logic as a Design Tool*

The use of common-emitter transistor amplifiers as an integral part of logic nets is commonly assumed to require NOT-AND (Sheffer Stroke) and/or NOT-OR (NOR) logic functions to represent the resultant logical behavior, as a result of the inverting properties of the common-emitter connected transistor. Logic design procedures based upon these functions have been described.[1,2]

A prevalent design approach is to derive an initial logic expression based on the elementary Boolean functions (AND, OR, and NOT) and then to perform a transformation into an expression based on inverting functions.

Example: Transform $Z=(A\cdot B)+(C\cdot D)+(E\cdot F)$ into an expression based on NOT-AND $(x\cdot y)'$ functions.

From the well-known Boolean transformation expression:

$$f(a+b+c+d\cdots)'=f(a'\cdot b'\cdot c'\cdot d'\cdots),$$

the expression can be rewritten as

$$Z'=(A\cdot B)'\cdot(C\cdot D)'\cdot(E\cdot F)'$$

and

$$Z=[(A\cdot B')\cdot(C\cdot D)'\cdot(E\cdot F)']'.$$

Compared with a conventional design based upon noninverting logic circuitry, the above procedure has several disadvantages:

1) The designer is forced to go through an additional algebraic manipulation, increasing both design time and chances of design errors.
2) It is difficult to visualize the behavior of the logic net from the transformed expression, compared with the original expression based on the elementary Boolean functions.

An alternative design method is described that permits the procedures for noninverting circuitry to be applied to inverting circuitry without additional algebraic manipulation. The method is essentially to regard the signal inversion taking place in the inverting circuitry to be a *physical* inversion rather than a *logical* inversion and to exclude it from the logical expression. The exclusion is accomplished by using a dual-polarity logic convention; that is, the polarity of the physical signal chosen to represent a logical assertion is changed between the input and the output of logic elements.

To illustrate the method, assume a circuit as that shown in Fig. 1. This circuit generates functions as follows, assuming a conventional single-logic polarity:

* Received by the PGEC, March 10, 1959.

[1] W. D. Rowe, "The transistor NOR circuit," 1957 WESCON Convention Record, pt. 4, pp. 231–245.

[2] N. T. Grisamore, L. S. Rotolo, and G. U. Uyehar, "Logical design using the stroke function," IRE Trans. on Electronic Computers, vol. EC-7, pp. 181–183; June, 1958.

Logic Polarity

Positive=Assert $C=(A\cdot B)'$ NOT-AND

Negative=Assert $C=(A+B)'$ NOT-OR

Using a dual-polarity logic, the following obtains

Input positive and output negative

=Assert $C=A\cdot B$ AND

Input negative and output positive

=Assert $C=A\cdot B$ OR.

Several problems associated with the use of a dual-polarity logic must be considered. The first is one of defining the assumed polarity symbolically. A proposed method is to use solid lines on the flow diagram to represent one polarity and dashed lines to represent the opposite polarity. Thus, in the example above, solid lines might represent positive=assert and dashed lines negative =assert. The AND and OR symbols then would become as shown in Fig. 2 (for the circuit of Fig. 1).

The second problem derives from the fact that logical negations and physical inversions are being differentiated. For example, if the circuit of Fig. 1 is used with a single input, the resultant behavior may be either a logical negation (NOT) or a physical inversion, which we choose to call an INVERT. In terms of the design procedure, the two are easily separated:

1) An element used in a NOT sense is assumed to have the same polarity convention on both input and output, and to negate the logic.
2) An element used in an INVERT sense is assumed to have opposite polarity conventions on input and output but not to modify the logic. An invert element is used to obtain a correct polarity for entry to AND/OR elements, when such is not at hand.

The symbology for the NOT and INVERT operations then are as shown in Fig. 3, based on the assumptions used in Fig. 2.

A third problem is the following the previously described procedure will occasionally lead to a redundant situation in the shape of a NOT element followed by an INVERT element. The difficulty is circumvented by adopted a symbolic operation that has been termed an INTERCHANGE. The INTERCHANGE symbol on the flow diagram directs that the logic polarity be reversed and the same time defines a NOT operation. Fig. 4 illustrates the removal of an INVERT-NOT redundancy through the use of a proposed INTERCHANGE symbol. The nature of the flow lines on each side of the symbol clearly indicates a polarity inversion; the logic negation may be left implied or the actual logic expression may be written on both sides of the symbol to assist the user of the flow diagram.

The design procedure based on the foregoing is as follows:

1) The system logical functions are evolved upon elementary Boolean expressions.

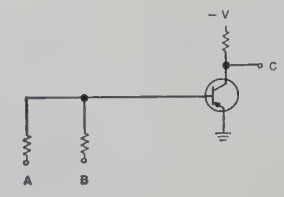


Fig. 1—Resistance-coupled transistor logic circuit using a p-n-p transistor.

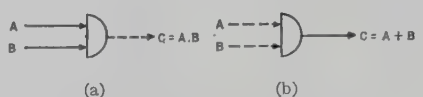


Fig. 2—Symbols for (a) AND and (b) OR incorporating polarity definitions.

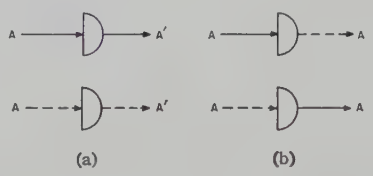


Fig. 3—Symbols for the (a) NOT and the (b) INVERT operations.

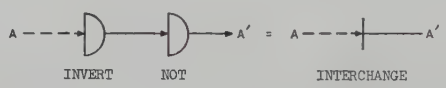


Fig. 4—Removal of a redundancy by use of the INTERCHANGE symbol.

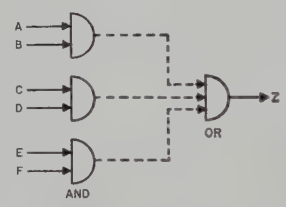


Fig. 5—Logic flow diagram for the expression $Z=A\cdot B+C\cdot D+E\cdot F$ using dual-polarity logic.

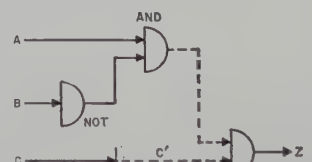


Fig. 6—Logic flow diagram for the expression $Z=(A\cdot B')+C'$.

2) The logic flow diagram is drawn using the polarity defining symbols shown in Fig. 2 and Fig. 3.
3) The INVERT element is inserted where the correct polarity is not at hand for entry to AND, OR elements.
4) The INTERCHANGE symbol is used to eliminate or avoid INVERT-NOT redundancies.

To illustrate the use of the method, Fig. 5 shows the logic flow diagram for the previously discussed expression $Z=(A\cdot B)+(C\cdot D)+(E\cdot F)$, using dual-polarity logic. An additional illustration is given in Fig. 6 for the expression $Z=(A\cdot B')+C$ where the use of the INTERCHANGE symbol is shown.

Experience has shown that this procedure simplifies the task of the logic designer when using inverting circuitry. It is true that the individuals dealing with the actual circuitry must consider two polarities for the logical assert condition rather than one; this disadvantage appears to be more than compensated by the fact that the system behavior is expressed for the most part in terms of the elementary and familiar Boolean symbols.

Paul M. Kintner
Airborne Instruments Lab.
Mineola, N. Y.

Analysis of Binary Time Series in Periodic Functions*

In this article it is shown that for binary time series (defined only for $t=0, 1, 2$, etc., and with values either 0 or 1), there is an analysis somewhat analogous to the Fourier analysis for functions of a continuous argument. The analogy is that the functions in which a binary time series can be analyzed are periodic. The difference with Fourier analysis is first that in our analysis the periods turn out to be n, $\frac{1}{2}n$, $\frac{1}{4}n$, $\frac{1}{8}n$, etc. (down to period 1) where n is the total length of the series, as compared to the periods n, $\frac{1}{2}n$, $\frac{1}{3}n$, etc. (down to period 0) in Fourier analysis. As a consequence, the analysis is only possible for a length $n=2^p$, while Fourier analysis, of course, is possible for an interval or arbitrary length. Furthermore, in the binary case the functions turn out to be not mutually orthogonal.

The periodic binary functions $f_k(t)$ are given by

$$f_k(t)=\frac{(t+k)!}{t!k!}=\binom{t+k}{k},$$

taken "modulo two." $f_1(t)$ comes out to be 1010 . . . , with period 2, while $f_2(t)$ is 11001100 . . . with period 4. f_0 will be defined as the series 1111 . . . . In general now it can be proved that $f_k(t)$ has a period 2^{p+1}, if $2^p\leqq k<2^{p+1}$. To prove this, we need some knowledge of the theory of delay polynomials, as given by Huffman,[1] who defines the operator D as applied to $f(t)$ so that $Df(t)=f(t-1)$, and $D^kf(t)=f(t-k)$. A delay polynomial $P(D)$ is now an operator

$$P(D)=C_0+C_1D+C_2D^2+\cdots C_kD^k,$$

$+$ standing for "+modulo two," as it does throughout this paper. So

$$P(D)f(t)=(C_0+C_1D+\cdots C_kD^k)f(t)$$
$$=C_0f(t)+C_1f(t-1)+\cdots C_kf(t-k).$$

* Received by the PGEC, February 9, 1959.

[1] D. A. Huffman, "The synthesis of linear sequential coding networks," *Proc. Third Symp. on Information Theory*, London, Eng., p. 77; 1955.

An important concept is the operator $P^{-1}(D)$. It is the operator which restores $f(t)$ from $P(D)f(t)$. Huffman shows how these operators can be realized by networks containing only two simple kinds of elements, one for adding and one for delaying one digit.

The proof for the period of $f_k(t)$ now goes as follows: First one can prove that

$$(1 + D)f_k(t) = f_{k-1}(t),$$

which follows in a straighforward manner from the definition of f_k:

$$\begin{aligned}(1 + D)f_k(t) &= (1 + D)\frac{(t + k)!}{t!k!} \\ &= \frac{(t + k)!}{t!k!} + \frac{(t + k - 1)!}{(t - 1)!k!} \\ &= \frac{(t + k - 1)!}{(t - 1)!k!}\left[1 + \frac{t + k}{t}\right] \\ &= \frac{(t + k - 1)!}{(t - 1)!k!}\left[1 + 1 + \frac{k}{t}\right] \\ &= \frac{(t + k - 1)!}{t!(k - 1)!} \\ &= f_{k-1}(t); \text{ (since } 1 + 1 = 0).\end{aligned}$$

Consequently, $(1+D)^k f_k(t) = f_0(t) = 1111 \cdots$. This allows us to construct

$$f_k(t) : f_k(t) = \frac{1}{(1 + D)^k} f_0(t).$$

If $k=2^p$, then $(1+D)^k = 1+D^k$. It follows now from the network construction of $1/(1+D^k)$ (the reader will have to consult Huffman's article[1] for this), that

$$\frac{1}{1 + D^k}\left[f_0 = 1111 \cdots\right]$$

is given by a periodic function of k ones followed by k zeros, etc. The period is therefore $2k = 2^{p+1}$. On the other hand, if $k = 2^p - 1$, $(1+D)^k$ is given by $(1+D+D^2+ \cdots +D^k)$. Again the network construction of

$$\frac{1}{1 + D + D^2 + \cdots D^k} f_0(t)$$

shows that this is a one followed by k zeros, etc. The period therefore is 2^p. In general a function f_{k-1} can have a period which is shorter than the one of f_k, but it cannot be longer, since $f_{k-1} = (1+D)f_k$. It follows that all functions f_k must have a period shorter than or equal to the period 2^p, if $k < 2^p$, and it must at least be 2^p, if $k \geqq 2^{p-1}$, and this completes the proof.

Next, we must prove that every function $f(t)$, defined for n digits, can be written as a sum of f_k's, provided $n = 2^p$:

$$f(t) = \sum_0^{n-1} i\, A_i f_i(t),$$

where the A_i's are the constant coefficients. The proof consists simply of a computation of the A_i's. Take

$$(1 + D)^k f(t = n - 1) = \sum_0^{n-1} i\, A_i (1 + D)^k f_i(n - 1).$$

It can be proved now that $(1+D)^k f_i(n-1) = 0$ for $k \neq i$, and so: $A_i = (1+D)^i f(n-1)$. To prove this, one has only to consider that $f_{i-k}(n-1) = 0$ unless $i-k=0$. A complete proof would require a demonstration that $f(t)$ is actually represented by $\sum A_i f_i$. This follows from the fact that the "Fourier coefficients" of $f(t) + \sum A_i f_i$ would all be zero, and that a function with all A's equal to zero is identical zero.

The harmonic analysis given above has the disadvantage that all A's depend on $f(n-1)$, the last value of the series. With only slight modification one can make a harmonic analysis in which a new coefficient, call it B_k, can be obtained for every new value $f(k)$ in the time series. The periodic functions $g_i(t)$ are now the time reverse of $f_i(t)$. For instance: $g_0(t) = 1111 \cdots$; $g_1(t) = 0101 \cdots$; $g_2(t) = 00110011 \cdots$. Then

$$f(t) = \sum_0^{n-1} i\, B_i g_i(t),$$

and

$$B_i = (1 + D)^i f(i).$$

The proof of this formula can be obtained by considering $f(t)$ as the time reverse of a function $f'(t)$ since

$$f'(t) = \sum_0^{n-1} i\, A_i' f_i(t),$$

$$B_k = A'_k = (1 + D)^k f'(n - 1),$$

and $B_0 = f'(n-1) = f(0)$; $B_1 = (1+D)f'(n-1) = (1+D)^1 f(1)$, and, in general, $B_k = (1+D)^k f(k)$. The analysis of $f(t)$ in the B_k's can also be considered as a transformation of $f(t)$ into another time series, by taking $B_t = B(t)$. This is a linear transformation, which can be given by the transformation matrix (α_n). For instance, for $n=4$ we have

$$\begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} B(0) \\ B(1) \\ B(2) \\ B(3) \end{bmatrix}.$$

This example brings out the general form of the matrix, which is such that

$$(\alpha_{2n}) = \begin{pmatrix} \alpha_n & 0 \\ \alpha_n & \alpha_n \end{pmatrix}.$$

The general proof for this formula will be omitted here, but from this it also follows that $(\alpha_n)^2 = (1)$. This means that a harmonic analysis of the time series $B(t)$ will give us the original time series $f(t)$ back.

Finally, one more difference with Fourier analysis is that the functions f_k, or g_k, are not mutually orthogonal, like Fourier components. A set of mutual orthogonal functions in the binary system, however, would be a rather trivial one. It would allow only a single 1 in each row or column of the transformation matrix. As a consequence of the nonorthogonality, one cannot compute the B_k's as an integral of $f(t)g_k(t)$ over the interval. There is however a slightly different formula:

$$B_k = \sum_0^{n-1} f_{n-k-1} f(t).$$

The reader can check this formula easiest from the matrix representation, but a proof will be omitted.

A final remark is that this analysis can be extended to functions of more variables, say x and y, by considering the functions

$$\frac{(x + k_1)!}{x!k_1!} \frac{(y + k_2)!}{y!k_2!}.$$

The author would like to thank Dr. J. Hartmanis for valuable criticism.

P. J. VAN HEERDEN
General Electric Res. Lab.
Schenectady, N. Y.

Residues of Binary Numbers Modulo Three*

A method is described for determining the remainder on dividing a binary number by three, without actually dividing out. It may be of interest for test routines or error-correcting codes (ternary instead of parity checks).

Rule 1—Where a pair of adjacent zeros or ones occur, cross them off. The digits immediately adjacent to a crossed-off pair are then considered adjacent to each other. Continue the crossing-off process as far as possible. The ultimate result is an alternating sequence of zeros and ones. Cross off an initial zero (if present) so that the sequence begins with one and terminates with zero or one.

Rule 2—Cross off digit groups containing three ones. If only zeroes remain, the original number is exactly divisible by three. If the least significant digit is one (respectively zero) and a single 1 is not crossed out, the remainder on division of the original number by three is one (respectively two), while if two ones are not crossed out the remainder is two (respectively one).

Proof—Replacing 11 by 00 subtracts a multiple of three. Crossing off 00 and considering its neighboring digits as adjacent replaces a number of the form $4a+b$ by $a+b$, which has the same residue modulo three. Rule 1 therefore yields a number with the same residue modulo three as the original number. Crossing off sequences containing three ones by Rule 2 subtracts a multiple of 10101, *i.e.*, 21, and thus does not change the remainder. The rest of the rule then says the remainder is one if the uncrossed out "tail" is 1 or 1010, two if the tail is 10 or 101, as is obvious.

Extension of the rule to numbers of the form 3×2^n is left as an exercise for the reader!

JEROME ROTHSTEIN
Edgerton, Germeshausen and Grier, Inc.
Boston 15, Mass.

* Received by the PGEC, April 13, 1959.

Comments on Optimum Character Recognition Systems*

Recently, this author came across Flores' correspondence[1] on the above paper.[2] I wish to thank him for his comments and would like to emphasize some points which are relevant to the statement with which he has taken issue.

On the Statement Under Discussion

The statement reads: "The above results also indicate that the recognition system described in Section II (correlation coefficient) is not optimum for the particular signal and noise structure as given in examples 1 or 2."[1,2] This statement was made within the framework of two general requirements, namely:

Requirement 1: Character recognition systems have an option to reject [page 249, above (4), reference 2].

Requirement 2: Consequences of various decisions are such that weights satisfy the following inequalities [(7) in reference 2]

$$w_{ij} > w_{i0} > w_{ii} \quad \text{for all } i \neq j \neq 0.$$

It seems to me that a provision for rejection is of utmost importance in character recognition systems. The inequalities mean that to misidentify a character is most costly and to reject is more costly than to identify correctly. The validity of the statement is, of course, questionable if the requirements are changed.

To illustrate the effect of change in requirements on the resultant optimum system, suppose the second requirement is modified to read: $w_{ij} > w_{i0} = w_{ii} (i \neq j \neq 0)$. *i.e.*, the consequences of rejection and correct recognition are the same. Such weights are used in Flores' example.[1] As a result, a system which makes rejections all the time would be optimum. It will not attempt to identify any character at the risk of misrecognition, when it incurs no more loss to reject than to correctly identify characters.

On Flores' Example

Flores[1] assigns that $w_{ii} = w_{i0} = 0$ for all i, and rejection is not considered. These two assumptions are contrary to the two requirements, and in consequence, the statement under discussion is not necessarily valid. I agree with Flores that, for certain noise statistics, the correlation system is optimum among systems having *no* option to reject.[3] However, in my opinion, a character recognition system with no provision for rejection is not "a most important case."

In reference 2, σ_{ij}^2 is the variance of noise of the jth component of the ith character. It is not a covariance, and, in general, $\sigma_{ij} \neq \sigma_{ji}$ and $\sigma_{ij} \neq 0$. There seems to be a misinterpretation when Flores took σ_{ij}^2 as covariance and concluded that $\sigma_{ij} = 0$ for $i \neq j$.

Correlation is Optimum Under Some Different Requirements

The correlation system may be optimum if requirements are different. For example, consider a case wherein rejection is not allowed. This is equivalent to assigning weights such that $w_{i0} \geq w_{ij}$ for all i and $j \neq i$, *i.e.*, misrecognition incurs no worse consequence than rejection. Let the noise statistics be as described in example 1 of reference 2 with an additional condition that σ_{ij} is independent of the character, *i.e.*, $\sigma_{ij} = \sigma_j$ for all i. Call

$$1/K = (2\pi)^{s/2} \prod_{j=1}^{s} \sigma_j.$$

Eq. (29) of reference 2 reduces to

$$F(v \mid a_i) = K \exp\left[- \sum_{j=1}^{s} \frac{(v_j - a_{ij})^2}{2\sigma_j^2} \right],$$

which is the same equation as appears in the middle of reference 1, although σ_j is defined differently. The rest of reference 1 essentially follows. To find the minimum $X_i(v)$ means to find maximum

$$\sum_{j=1}^{s} (a_{ij} v_j / \sigma_j)^2 - \frac{1}{2} \sum_{j=1}^{s} (a_{ij}^2 / \sigma_j^2).$$

This is the last equation in reference 1. (There seem to be some typographical errors.)

Thus, the correlation system with bias term for waveform power is optimum, if occurrence probabilities are equal (*i.e.*, $p_i = 1/c$ for all i), if weights are given as: $w_{ij} = w_m$ for $i \neq j$ and $w_{ii} = w_c < w_m$; and if rejection is not allowed.

C. K. Chow
Burroughs Corporation Res. Center
Paoli, Pa.

Solution to Boolean Equations*

I would like to point out that the paper by Rouche[1] is concerned with a problem that has already been solved in the literature.[2,3]

The solutions to Boolean equations, as developed in the cited papers, were not applied to the logical design of digital circuitry. However, it may be of interest to note that a great deal of work along these lines has been accomplished at the National Bureau of Standards since 1955. Some of the results of this work are presently being prepared and submitted for publication; they will also appear in the textbook, "Digital Computers and Control Engineering," in preparation for McGraw-Hill Book Company, Inc.

To relate the two referenced papers to the work of Rouche, it can be observed that Rouche's matrix B is precisely the permutation R^{-1} introduced in the earlier paper,[2] and is a special case of the (R_{ji}) matrix introduced in the later one.[3] Rouche's problem appears as a special case of the "change of variables" problem[2] or as a special case of the antecedence and consequence problem.[3]

Robert S. Ledley
National Bureau of Standards
Washington 25, D. C.

* Received by the PGEC, September 15, 1958.

[1] I. Flores, "An optimum character recognition system using decision functions," IRE Trans. on Electronic Computers, vol. EC-7, p. 180; June, 1958.

[2] C. K. Chow, "An optimum character recognition system using decision functions," IRE Trans. on Electronic Computers, vol. EC-6, pp. 247–254; December, 1957.

[3] An example is given by C. K. Chow, in "An Optimum System for Character Recognition Using Decision Functions," Burroughs Res. Center Tech. Rept., TR56-78; December 6, 1956.

* Received by the PGEC, March 20, 1959.

[1] N. Rouche, "Some properties of Boolean equations," IRE Trans. on Electronic Computers, vol. EC-7, pp. 291–298; December, 1958.

[2] R. S. Ledley, "Mathematical foundations and computational methods for a digital logic machine," *J. Operations Res. Soc. Am.*, vol. 2, pp. 249–274; August, 1954.

[3] R. S. Ledley, "Digital computational methods in symbolic logic, with examples in biochemistry," *Proc. Natl. Acad. Sci.*, vol. 41, pp. 498–511; July, 1955.

Contributors

Robert B. Ash (M'58) was born in New York, N. Y., on May 20, 1935. He received the B.A. degree in 1955, and the B.S. and M.S. degrees in electrical engineering in 1956 and 1957, respectively, from Columbia University, New York, N. Y.

R. B. Ash

During the summer of 1957 he did research in reliability theory at the IBM Watson Laboratory. He is now an instructor in electrical engineering and a candidate for the Ph.D. degree in electrical engineering at Columbia University. He is engaged in research on network topology and information metworks.

Mr. Ash is a member of Phi Beta Kappa, Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

❖

George A. Bekey (M'56) was born in Czechoslovakia, on June 19, 1928. He received the B.S. degree in electrical engineering from the University of California, Berkeley, in 1950, and the M.S. degree in engineering from the University of California at Los Angeles, in 1952.

G. A. Bekey

His experience includes three years as a research engineer in the Department of Engineering, U.C.L.A., working with analog computing equipment including the ac network analyzer and the mechanical and electronic differential analyzers. Following service in the Signal Corps, he joined the Berkeley division of Beckman Instruments in October, 1955, as chief computer applications' engineer. From January, 1957, to June, 1958, he directed the activities of the Beckman Computation Center in Los Angeles, which he established.

In June, 1958, he joined the technical staff of Space Technology Laboratories, where he is concerned with analysis and computer simulation of control problems.

Mr. Bekey is a member of the AIEE and ACM.

❖

Eric E. Bittmann (M'54) was born in St. Moritz, Switzerland, in October 1923. He received a diploma in electrical engineering from the Swiss Federal Institute of Technology, Zurich, Switzerland.

From 1948 to 1951, as a senior engineer for the Philco Corporation, he helped to develop color television transmitters, flying-spot scanners and monitors. From 1951 to 1955 he was involved with the project of telemetry ground stations for the Applied Science Corporation of Princeton, N. J.

E. E. Bittmann

He joined the Burroughs Corporation Research Center early in 1955 and was engaged in work on large coincident current core memories.

Since late 1957, as project engineer, he has been concerned with the development of thin film memories.

Mr. Bittmann is a member of RESA.

❖

Gerrit A. Blaauw (S'51–A'53–M'57) was born at The Hague, Netherlands, on July 17, 1924. He studied at the Institute of Technology at Delft, Netherlands. He received the B.S. degree in electrical engineering from Lafayette College in 1948, and the Ph.D. degree in applied science from Harvard University in 1952.

G. A. Blaauw

While at Harvard, he was a member of the staff of the Computation Laboratory and participated in the design of the Mark III and Mark IV calculators. From 1952 to 1955, he was a member of the staff of the Mathematics Center, Amsterdam, Netherlands, where he cooperated in the design of of the ARRA and FERTA computers.

In May, 1955, he joined the IBM Product Development Laboratory, Poughkeepsie, N. Y. He has been engaged in the systems planning of various machine projects.

Dr. Blaauw is a member of the Association for Computing Machinery and of Sigma Xi.

❖

Erich Bloch (A'52–M'57) was born in Sulzburg, Germany, on January 9, 1925. He attended the Federal Polytechnic Institute, Zurich, Switzerland and received the B.S.E.E. degree in 1952 from the University of Buffalo, N. Y.

He joined the International Business Machines Corporation in 1952 and worked on several aspects of computer development and design. He specialized in magnetic core storage units, and was in charge of a core logic group at the IBM Research Center. Mr. Bloch is now a senior engineer in charge of computer systems design, in the IBM Los Alamos computer project.

E. Bloch

❖

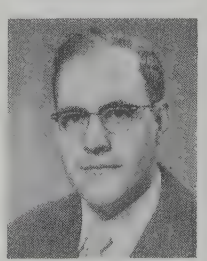
Thomas E. Bray (S'53–A'55) was born on September 12, 1932, in Chicago, Ill. On receiving the B.S. degree in electrical engineering from the University of Wisconsin, Madison, in 1954, he joined the General Electric Company. He completed the Advanced Technical Course and and the Electronics "C" Course of the Advanced Engineering Program, with assignments in radar systems, transistor development, and studies of ferroelectric and electro-optical circuits.

T. E. Bray

After one year with the Advanced Engineering Program he joined the Advanced Components and Networks Component of Electronics Laboratory, in 1958, engaging in development and studies of electro-optical circuits, magnetic switching circuits, and optical detectors.

❖

Frederick P. Brooks, Jr. (S'54–M'57) was born in Durham, N. C., in 1931. He received the A.B. degree in physics from Duke University in 1953. He did graduate work in the design and application of computing systems at Harvard University, receiving the S.M. and Ph.D. degrees in 1955 and 1956, respectively.



F. P. Brooks, Jr.

Since 1956, he has been with the IBM Product Development Laboratory, Poughkeepsie, N. Y., where he is at present an advisory engineer in the group planning the IBM Stretch computer.

Dr. Brooks is a member of the Association for Computing Machinery, Phi Beta Kappa, and Sigma Xi.

Werner Buchholz (S'44–A'49–M'52–SM'54) was born in Germany on October 24, 1922. He received the B.A.S. degree in 1945 and the M.A.S. degree in 1947 from the University of Toronto. In 1946, he went to the California Institute of Technology as a graduate student and instructor. While there, he participated in the development of the Caltech Analog Computer and, in 1950 obtained the Ph.D. degree in electrical engineering.

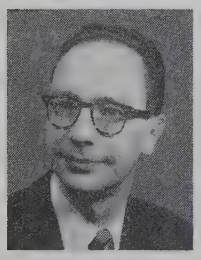
W. Buchholz

In 1949, Dr. Buchholz joined the IBM Laboratories in Poughkeepsie, N. Y. He participated in a number of systems planning projects in research and product development, including the planning of the IBM 701 and 702 computers. He is currently manager of engineering planning on IBM's Stretch computer.

He is a member of the Association for Computing Machinery.

❖

Leon R. Bush was born in New York, N. Y., on April 14, 1929. He received the Bachelor of Engineering Physics degree from Cornell University, Ithaca, N. Y., in 1951, and has taken graduate work in mathematics and physics at the University of Buffalo, N. Y.

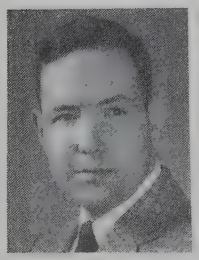


L. R. Bush

Mr. Bush has been working in the field of systems analysis and guided missile systems at the Cornell Aeronautical Laboratory, Inc., Buffalo, since 1951. At present, he is head of the analysis section of the Weapon Systems Design Department.

❖

Fred B. Cox (S'58) was born in Dallas, Texas, on June 3, 1934. He received the B.S.E.E. degree from the Agricultural and Mechanical College of Texas in 1956 and the S.M.E.E. degree from the Massachusetts Institute of Technology in 1958, where he is currently working toward the Electrical Engineer's degree.



F. B. Cox

At Texas A. and M., he worked in the Network Calculator Laboratory on the simulation of electrical power systems. He has been associated with Convair, Fort Worth, and with the Sperry Gyroscope Company. Since 1957 he has been a research assistant in the M.I.T. Servomechanisms Laboratory, working on digital and analog-digital techniques for simulation of physical systems.

Mr. Cox is a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, and Phi Eta Sigma.

❖

Harvey L. Garner (S'51–A'51–M'57) was born on December 23, 1926, at Lake, Colo. He received the B.S. and M.S. degrees from the University of Denver, and the Ph.D. degree from the University of Michigan.

H. L. Garner

From 1949 to 1951 he participated in the cosmic ray research program at the University of Denver and the Inter-University High-Altitude Research Laboratory at Echo Lake, Colo. From 1951 to 1955 he was associated with the development and operation of the MIDSAC and MIDAC computers at the Engineering Research Institute of the University of Michigan. In 1955 he became an instructor in the Department of Electrical Engineering. In this capacity he has been active as a co-supervisor of the MIC computer project and has been in charge of several special intensive summer computer courses.

Dr. Garner is presently assistant professor of electrical engineering at the University of Michigan and also serves as consultant to several electronic firms. He is a member of Sigma Xi, Sigma Pi Sigma, and the Association for Computing Machinery.

❖

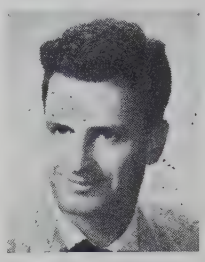
Elmer G. Gilbert (S'51–A'52–M'57) was born on March 29, 1930, in Joliet, Ill. He received the B.S.E. degree in electrical engineering in 1952, the M.S.E. degree in electrical engineering in 1953, and the Ph.D. degree in instrumentation engineering in 1957, all from the University of Michigan.

E. G. Gilbert

Since 1953 he has taught in the Department of Aeronautical Engineering at the University of Michigan. Presently, he is an assistant professor of aeronautical engineering and teaches courses for the instrumentation engineering degree program. His research activities included work in automatic control, electronic analog computation, and related fields.

Dr. Gilbert is a member of Eta Kappa Nu, Phi Kappa Phi, Tau Beta Pi, and Sigma Xi.

Victor H. Grinich was born in Aberdeen, Wash., on November 26, 1924. He received the B.S. degree in electrical engineering from the University of Washington in 1945 and the M.S.E.E. degree from the same University in 1950. He received the Ph.D. degree in electrical engineering from Stanford University in 1953.



V. H. Grinich

From 1946 to 1948 he taught in the Electrical Engineering Department at the University of Washington. From 1952 to 1956 he was with the Stanford Research Institute in connection with the applications of transistors to color television. He joined the Shockley Semiconductor Laboratory in 1956, engaged in instrumentation and silicon device applications. He has been with Fairchild Semiconductor Corporation since 1957, as Head of Application Engineering and Device Evaluation, working with silicon transistors and other semiconductor devices.

❖

Isy Haas was born in Istanbul, Turkey, on June 18, 1934. He received the B.S. degree in electrical engineering from the Robert College in Istanbul in 1955 and the M.S.E. degree from Princeton University in 1957.

I. Haas

He spent the summer of 1956 with the IBM Research Laboratories, Poughkeepsie, N. Y., engaged in research on superconductivity. He joined Remington Rand Univac in 1957, working on transistor circuit design for the LARC computer. He has been with Fairchild Semiconductor Corporation since July, 1958, engaged in research in circuitry for semiconductor devices.

❖

Ralph R. Hartel (M'52) was born in Quincy, Mass., on April 8, 1928. From 1946 to 1948, he served in the U. S. Navy as Division Petty Officer, Electronic, *U.S.S. Springfield*, and at the Naval Research Laboratory.

R. R. Hartel

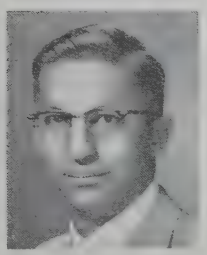
In 1948, he joined RCA and became, successively, chief field engineer in the 6th Naval District, instrumentation project engineer for RCA at the U. S. Navy

Department in Washington, manager of RCA's missile tracking station in the West Indies, and while a member of the chief engineer's technical staff, systems engineering, at Cocoa, Fla., was acting RCA project engineer for data collection on the U. S. Army's Redstone missile.

He joined Lake Service Corp., Boston, Mass., as chief engineer, in 1955. He went to The National Cash Register Company, Electronics Division, in 1957 as a senior research engineer and worked on digital storage using magnetic recording. In 1958, he joined Clevite Electronic Corp., Cleveland, Ohio, as chief project engineer for magnetics and he is now manager of the Magnetic Recording Department.

❖

Robert S. Ledley (M'58) was born on June 28, 1926, in New York, N. Y. He received the D.D.S. degree from New York University in 1948, and the M.A. degree in mathematical physics from Columbia University, New York, N. Y., in 1949. He joined the National Bureau of Standards, Washington, D. C., in 1951, working in biophysics and the logical design of digital computers. In 1954, he became a staff member of the Operations Research Office, where he worked on computer simulations and the application of logic to operations research problems.



R. S. Ledley

Since 1956, he has been an assistant professor of electrical engineering at George Washington University, Washington, D. C. He is also a staff member of the National Academy of Sciences-National Research Council project on the use of computers in biomedical sciences. He is consultant mathematician at the National Bureau of Standards, data processing systems division, and consultant to the National Library of Medicine. Currently, he is principal investigator of several NIH and NSF grants on information retrieval systems, computational methods in symbolic logic, and computers in biomedical research.

Dr. Ledley is a member of the American Mathematical Society, American Physical Society, Operations Research Society of America, and the American Museum of Natural History.

❖

Richard C. Lee (S'55–M'57) was born in Atlanta, Ga. on October 5, 1934. He received the B.S.E.E. degree from Duke University in 1956 and the S.M.E.E. degree from the Massachusetts Institute of Technology in 1958, where he is currently working toward the electrical engineer's degree.

He has been associated with RCA and Autonetics, a division of North American Aviation, Inc. Since 1956, he has been a research assistant in the M.I.T. Servomechanisms Laboratory, working on combined analog-digital techniques for the simulation of aircraft in real time.

R. C. Lee

Mr. Lee is a member of Eta Kappa Nu, Sigma Xi, Phi Beta Kappa, Tau Beta Pi, and the Association for Computing Machinery.

❖

Petar Madich was born on June 25, 1922 at Rashka, Yugoslavia. He graduated from the Electrical Engineering Faculty of the University of Belgrade in 1949.

His first employment was with the Institute of Electrical Power in Belgrade where he studied the problems of transmission of electrical power with the aid of a network analyzer. In 1951 he joined the Boris Kidrick Institute of Nuclear Sciences in Belgrade and took part in the construction of an analog computer for solving systems of linear algebraic equations and a repetitive differential analyzer. Later on he studied the problems of ill-conditioned systems of linear algebraic equations. He is now working on the repetitive differential analyzer and is chief of the Department of Numerical Analysis at the Boris Kidrich Institute.

P. Madich

❖

Sanford M. Marcus (M'58) was born in New York, N. Y. on March 18, 1932. He received the Bachelor of Science degree in physics from Brooklyn College, Brooklyn, N. Y. in 1954.

He then attended Columbia University, New York, N. Y., receiving the Master of Science degree in 1957.

Since 1957 Mr. Marcus has been associated with Radio Corporation of America, Camden, N. J., where he has been engaged in the development and application of semiconductor devices.

S. M. Marcus

John J. Miyata (A'52–M'57) was born in El Centro, Calif., on December 7, 1921. He received the B.S. degree in applied physics from the University of California at Los Angeles in 1949.

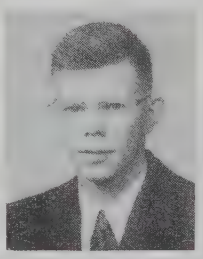
J. J. Miyata

From 1950 to 1952, he was a physicist on the guidance systems for the Sidewinder guided missile at the U. S. Naval Ordnance Test Station, China Lake, Calif. From 1952 to 1956, he worked at Northrop Aircraft Inc., Hawthorne, Calif., as a research engineer on the guidance system for the Snark missile. In 1956, he joined the Electronics Division of The National Cash Register Company, Hawthorne, Calif., where he is presently project engineer in the Research Department. He is responsible for directing a program of research and development in the field of magnetic recording for digital computers.

Mr. Miyata is a member of the American Physical Society.

❖

Walter L. Morgan (S'52–A'55) was born in Nutley, N. J., on December 20, 1930. He received the B.S. degree in electrical engineering from the Carnegie Institute of Technology, Pittsburgh, Pa., in 1954.



W. L. Morgan

From 1949 to 1954 he was associated with the Liberty Mirror Division of the Libbey Owens Ford Glass Company and worked on high-vacuum thermal evaporation methods of producing thin-film coatings. In 1954, he joined the Radio Corporation of America in Princeton and Moorestown, N. J., where he has been engaged in research and development work on digital data-handling devices. He is currently engaged in development work on magnetic flux-logic and memory devices.

Mr. Morgan is a member of the AIEE and Sigma Nu.

❖

John K. Munson was born in Minneapolis, Minn., June 11, 1930. He received the B.S. degree in chemical engineering from the University of Minnesota in 1951 and the M.S. in chemical engineering from the University of Illinois in 1953.

He was stationed at Redstone Arsenal for three years at the Ordnance Guided Mis-

sile School serving as a Nike missile system instructor. Since 1956 he has been with the Engineering Department of E. I. Du Pont de Nemours & Company, in the Instrumentation and Engineering Computation Sections, working on analog computer simulation of chemical process problems.

J. K. Munson

Mr. Munson is a member of Tau Beta Pi, Sigma Xi, and the American Institute of Chemical Engineers.

❖

Patrick J. Orlando was born in Weehawken, N. J., on July 11, 1921. He received the B.M.E. degree from Cooper Union, New York, N. Y., in 1942, and the M.S.M.E. degree from Stevens Institute of Technology, Hoboken, N. J., in 1950. He has taken graduate work in physics and mathematics at Columbia University, New York, N. Y., and with Cornell Extension Studies, Buffalo, N. Y.

P. J. Orlando

Mr. Orlando served with the U.S.A.F. as an aircraft maintenance officer from 1942 to 1946 and as a mechanical engineer from 1951 to 1953.

From 1946 to 1948, he was a development engineer at the Celanese Corporation, Newark, N. J., and from 1949 to 1951, a research engineer with the Worthington Corporation, Harrison, N. J.

Since 1953, Mr. Orlando has been with the Cornell Aeronautical Laboratory, Inc., Buffalo. He is now a principal physicist in the dynamic analysis branch of the Aeromechanics Department, where he is engaged in research in guided missile systems, computers, and applied mathematics.

❖

Nedeljko Parezanovich was born on August 25, 1932 at Ivanjica, Yugoslavia. He graduated from the Faculty of Natural Sciences of the University of Belgrade in 1957.

N. Parezanovich

He is now employed in the Boris Kidrich Institute of Nuclear Sciences in Belgrade and works on the repetitive differential analyzer in the Department of Numerical Analysis.

Robert C. Paulsen was born on January 18, 1919, in Parsippany, N. J. He studied electrical engineering at the Newark College of Engineering, New Jersey. During World War II he served for two years with the U. S. Army Signal Corps.

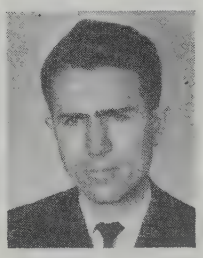
R. C. Paulsen

Mr. Paulsen has been with International Business Machines Corporation since 1940. His recent assignments include work in the applied core logic group of the IBM Research Laboratory. He is currently a staff engineer engaged in solid-state circuit work on the Los Alamos Computer project.

❖

Jovan Petrich was born on August 15, 1930 at Plevlje, Yugoslavia.

He graduated from the Faculty of Natural Sciences of the University of Belgrade in 1957.



J. Petrich

At the Boris Kidrich Institute of Nuclear Sciences in Belgrade, Mr. Petrich is working on the repetitive differential analyzer in the Department of Numerical Analysis.

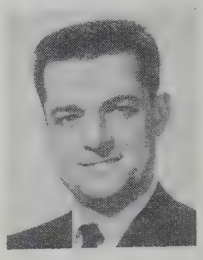
❖

Paul E. Pfeiffer was born in Newark, Ohio, in 1917, and received his public schooling in Houston, Texas. In 1938 he received the B.S. degree in electrical engineering from the Rice Institute. After two years of graduate work at Rice, he entered the School of Theology, Southern Methodist University, where he received the B.D. degree in 1943. After four years in the pastoral ministry, he returned to Rice to complete his graduate work and serve on the teaching staff. He received the M.S. degree in electrical engineering in 1948 and the Ph.D. degree in mathematics in 1952. At present, he is an associate professor of electrical engineering at Rice.

P. E. Pfeiffer

Dr. Pfeiffer is a member of the AIEE, the American Mathematical Society, the Association for Computing Machinery, the Mathematical Association of America, Sigma Xi, and Tau Beta Pi.

Robert V. Powell (SM'53–M'58) was born August 17, 1928, in Toledo, Ohio. Following five years as an electronic technician and instructor in the U. S. Navy, he attended the California Institute of Technology, where he received the B.S. and M.S. degrees in electrical engineering.



R. V. Powell

In 1956, his professional experience in the computing field began at Computer Engineering Associates in Pasadena, Calif., where he was engaged in engineering analysis on the direct analog computer until 1958. Early in 1958, Mr. Powell joined the computer staff at the Jet Propulsion Laboratory, where he is presently engaged in computer systems research and development.

❖

Donald E. Rosenheim (SM'56) was born on March 23, 1926, in New York, N. Y. His education was interrupted by service in the Navy, where he attended the Radio Materiel School and served two years as an electronic technician.

D. E. Rosenheim

He received the B.S. degree in electrical engineering magna cum laude from Polytechnic Institute of Brooklyn, N. Y., in 1949, and the degree of M.S. in electrical engineering from Columbia University, New York, N. Y., in 1957. From 1949 to 1951 he did development work on "servo-synchronized" transmitting equipment for the Servo Corporation of America.

In 1951 he joined the engineering laboratories of IBM Corporation at Poughkeepsie, N. Y. From 1951 to 1953 he was concerned with development work and instruction on the IBM 701 digital computer. Since 1952 he has been at the IBM Watson Laboratory at Columbia University, where he has been engaged in digital computer research high-speed logic and pulse techniques, and reliability research.

Mr. Rosenheim was a lecturer in electrical engineering at the College of the City of New York in 1956 and a lecturer in electrical engineering at Columbia University in 1958.

He is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

❖

Arthur I. Rubin (M'57) studied physics at the College of the City of New York, Columbia University, and Stevens Institute of Technology. He received the B.S. degree

from C.C.N.Y. in 1949, and the M.S. degree from Stevens Institute in 1953. In 1949 he joined the Research and Development Laboratory of the Picatinny Arsenal. There he was intimately associated with a wide variety of research and development projects in propellants and explosives. This work included high-pressure, high-temperature phenomena, gas kinetics, reaction rates theory, heat and thermodynamics, and the application of analog computers to problems in these areas.

A. I. Rubin

Since 1955, he has been with Electronic Associates, Inc., as an applications engineer, and in this capacity has been responsible for a wide variety of computer study projects in the fields of heat transfer, ballistic devices, guided missiles, fire control systems, nuclear reactor control and linear programming. Since 1957, he has been supervisor of analog computations for the Princeton Computation Center of Electronic Associates, Inc.

Mr. Rubin is a member of Phi Beta Kappa and the American Physical Society.

❖

J. Torkel Wallmark (A'48–M'55–SM'59) was born in Stockholm, Sweden, on June 4, 1919. He received the degree of Civilingenjor in electrical engineering in 1944, Teknologie Licentiat in 1947, and Teknologie Doktor in 1953 from the Royal Institute of Technology, Stockholm.

From 1944 to 1945 he was a vacuum tube designer with the A. B. Standard Radiofabrik. From 1945 to 1953 he was with the Royal Institute of Technology as a research assistant on vacuum tube problems, while spending periods at RCA Laboratories, Princeton, N. J., Elektrovarmeinstitutet, and Tekniska Forskningsradet, Stockholm, engaged in work on secondary emission tubes, semiconductors, and research administration, respectively. Since 1953 he has been with RCA Laboratories, Princeton, N. J., where he has been involved in magnetrons, color television, and lately, semiconductor devices.

J. T. Wallmark

Abstracts of Current Computer Literature

(THROUGH JANUARY, 1959)

This issue of the TRANSACTIONS continues the literature abstracting service recently inaugurated by the PGEC. The abstracts and associated subject and author indexes were prepared on a commercial basis by a Massachusetts firm under the management of Geoffrey Knight, Jr. This firm is best known for publication of the abstract journal, *Semiconductor Electronics*. Local volunteer support for this endeavor has been furnished by F. E. Heart, P. R. Bagley, and R. P. Mayer of the Lincoln Laboratory, M.I.T.

Comments on the abstract service are welcome. Our intent is to make it of greatest possible usefulness, consistent with the space and funds available, by suitable selectivity of items and clarity of abstract text. Can you suggest improvements? Additional copies of these abstracts are available from IRE Headquarters, 1 E. 79th St., New York 21, N. Y., at $1.00 per copy, or $3.50 for the set of four that are planned to appear in 1959.

Our attention has been called to the monthly computer bibliography published by Technical Information Company, Chancery House, London WC2, England. Their coverage is of a substantially larger number of items but with much briefer abstracts.—*The Editor*

TABLE OF CONTENTS

A-1: EQUIPMENT—THEORETICAL DESIGN

127
Logical Design for Digital Computers, by R. V. Norton (Western Elec. Co.); *Western Elec. Engr.*, vol. 11, pp. 28–36; October 1958.

A hypothetical stored program, general purpose, electronic digital computer is described as an illustration of some basic principles of computer logic design and operation. The basic circuits, arithmetic element, memory device, control element, and input-output equipment required by such a computer are discussed in some detail.

128
Computer for Solving Integral Formulations of Engineering Problems by Method of Successive Approximations, by J. M. Ham (Mass. Inst. Tech.); *U. S. Gov. Res. Repts.*, vol. 30, p. 385 (A); November 14, 1958. PB 126993.

It is concluded that small and relatively simple special-purpose computers designed for use by the research worker can contribute significantly to the effective application of methods of successive approximations. The design and application of such a computer is described. The machine is effective for evaluating integral transformations such as Fourier, correlation, and convolution integrals.

129
Relation Between Turing Machines and Actual Computing Machines, by J. B. Rosser (Cornell Univ.); *U. S. Gov. Res. Repts.*, vol. 30, p. 580 (A); December 12, 1958. PB 134543.

The idealized type of computing machine, known as a Turing machine, on which theoretical studies of computing machinery are usually based, is compared with actual machines. It is concluded that this sort of theoretical study is justified.

130
On the Minimum Logical Complexity Required for a General Purpose Computer, by S. P. Frankel; IRE Trans. on Electronic Computers, vol. EC-7, pp. 282–285; December, 1958.

A definition for the term "general purpose computer" (gpc) which is compatible with usage and is analogous to, but distinct from, Turing's definition of a "universal computer" is given. A gpc is presented in functional and logical design which seems to approximate the minimum complexity consistent with this definition. No specific definition of complexity nor a bound on required complexity is presented.

131
Analysis of Carry Transmission in Computer Addition, by S. G. Campbell and G. H. Rosser, Jr. (Duke Univ.); *U. S. Gov. Res. Repts.*, vol. 30, p. 392 (A); November 14, 1958. PB 132604.

A method of transforming immediately any information about ordinary binary carry into equivalent information about any transformed circuit is discussed. The analysis of carry transformations is valuable in approaching the general problem of carry transmission in computer addition. Characteristics of some carry transforming circuits are tabulated.

132
Logic Synthesis of Some High Speed Digital Comparators, by M. Nesenbergs and V. O. Mowery (Bell Tel. Labs.); *Bell Sys. Tech. J.*, vol. 38, pp. 19–44; January 1959.

Logical schemes for realizing high-speed digital comparators are derived by Boolean algebra methods. The requirements for speed and precision place serious restrictions on the switching circuits. In particular, the precision requirement makes direct subtraction by the use of analog devices undesirable; the speed requirement dictates that any carry structure should propagate from the most significant digit toward the least significant digits. Circuits satisfying the synthesis requirements and giving the sign and exact magnitude of the difference are derived first. These schemes are then modified and simplified to give the sign and approximate magnitude. Circuits giving only the sign of the difference are also derived.

A-2: EQUIPMENT—COMPONENTS AND CIRCUITS

133
Millimicrosecond Diffused Silicon Computer Diodes, by J. H. Forster and P. Zuk (Bell Tel. Labs.); 1958 WESCON Convention Record, pt. 3, pp. 122–130.

A diffused silicon computer diode with a switching time of about 1 mμs is described. Other characteristics favoring use in high-speed switching applications are low capacitance and forward dynamic resistance. DC reverse current is less than 0.015 μa at 20 volts, and voltage breakdown is in excess of 40 volts. Solid-state diffusion techniques are used both for reproducible junction formation and for precise introduction of the recombination centers which determine transient response. Low transition region capacitance is obtained with a small area graded junction. Other fabrication techniques permit high-temperature operation. Present data indicate that these diodes have a mean life (with reference to all test limits) greater than 100,000 hours when aged at 50 ma rectified dc. This life estimate (based on an assumption of an exponential failure rate) is increasing as data accumulate.

134
Millimicrosecond Magnetic Gating and Storage Element, by D. A. Meier (Natl. Cash Register Co.); *J. Appl. Phys.*, vol. 30, pp. 122–123 (L); January, 1959.

A new high-speed magnetic element for performing the logical gating and storage functions required in a digital computer is described. One configuration of the element consists of a round glass rod which is first chemically plated with a silver conductor and then electrodeposited with a Fe-Ni film. This device, called a ROD because of its shape, has an extremely square saturation loop, and the electroplated material exhibits a field threshold in which no flux change occurs. These properties make the ROD suitable as the storage element in a coincident-current selected memory and as a logical gate.

135
The Laddic—A Magnetic Device for Performing Logic, by U. F. Gianola and T. H. Crowley (Bell Tel. Labs.); *Bell Sys. Tech. J.*, vol. 38, pp. 45–72; January, 1959.

The Laddic, a ladder-like structure cut out of a ferrite having a rectangular hysteresis loop, is described. The sides of the ladder and all of the rungs are equal in minimum cross section so that all possible paths are flux-limited. The structure presents a large number of possible flux paths. By controlling the actual switching path through the structure, any Boolean function of n variables can be produced. A number of methods of operation are discussed, and design formulas and experimental results are presented. Switching speeds of a few tenths of a microsecond and repetition rates of a few hundred kilocycles have been achieved.

136
A High-Speed Logic System Using Magnetic Elements and Connecting Wire Only, by H. D. Crane (Stanford Res. Inst.); Proc. IRE, vol. 47, pp. 63–73; January, 1959.

Multiaperture magnetic elements with effective decoupling between different windings linking a given element are described. It is shown that by proper interconnection such elements can be used to provide unilateral information-flow properties without the use of explicit unilateral devices such as diodes. The resulting circuits require only magnetic elements and connecting wires and are relatively fast and inexpensive. Single-turn windings may be used in the coupling loops. Furthermore, inherent nondestructive read properties allow a great deal of logic facility. Simple shift register structures and logic elements and circuits are described.

137
Coincident Current Applications of Ferrite Apertured Plates, by W. G. Rumble and C. S. Warren (RCA); 1958 WESCON Convention Record, pt. 4, pp. 62–65.

A method of coincident current operation of ferrite apertured plates which requires only one hole per bit and a single digit winding for both reading and writing is described. Excellent results were obtained in experimental memories operated by economical transistor circuits and tested under adverse conditions. The plates are particularly suitable for making memory modules with which very compact memories of various sizes can be made.

138
Research and Development on Magnetic Films and Magnetic Matrix Memory Units, by E. R. Olson (Servomechanisms, Inc.); *U. S. Gov. Res. Repts.*, vol. 31, pp. 31–32 (A); January 16, 1959. PB 135053.

The final report for a research and development contract on evaporated magnetic films and magnetic memory matrices is presented. Methods and equipment for the deposition of thin films, fabrication of laminated elements, and testing of these elements are described. The deposition techniques studied include electron bombardment, direct evaporation, and spin coating of silicone resins. Nickel-iron alloy films were shown to exhibit certain desirable characteristics, and memory units containing nickel-iron were found to exhibit memory response characteristics.

139
Ferroelectric Storage Devices, by S. Morleigh (Powers-Samas Accounting Machines Ltd.); *Electronic Engrg.*, vol. 30, pp. 678–684; December, 1958.

The use of ferroelectric materials as storage devices in digital computers is discussed. A careful examination is made of switching characteristics and other relevant factors, including the limitation imposed by the present state of development of these materials. The storage properties of barium titanate single crystal capacitors are emphasized.

140
A Basic Transistor Circuit for the Construction of Digital-Computing Systems, by P. L. Cloot (Metropolitan-Vickers Elec. Co., Ltd.); *Metropolitan-Vickers Gazette*, vol. 29, pp. 298–306; November, 1958.

A basic circuit using one transistor, one capacitor, and three resistors, from which a complete digital-computing system may be economically constructed, is described. A system using this circuit is extremely simple to design, construct, and maintain, and should prove very reliable although it cannot achieve the speed of operation of systems using more complex circuits.

141
Bistable Circuits Using Unijunction Transistors, by T. P. Sylvan (G.E. Co.); *Electronics*, vol. 31, pp. 89–91; December 19, 1958.

The use of silicon unijunction transistors to simplify bistable circuit design and to permit operation at high ambient temperatures is described. Three bistable circuits, each of which utilizes the negative resistance characteristic of a single unijunction transistor, are shown. A ring counter based on one of these circuits operates at frequencies from 0 to 40 kc and at ambient temperatures up to 110°C.

A-3: EQUIPMENT—SUBSYSTEMS

142
A Magnetic Core Parallel Adder, by M. C. Chen (Stanford Univ.); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 262–264; December, 1958.

A logical design of a parallel adder using magnetic core elements which does not have the usual carry time limitations is described. The synthesis uses a truth-table technique.

143
Transistorized Decade Counter, by A. Szerlip (Packard Bell Electronics Corp.); 1958 WESCON CONVENTION RECORD, pt. 6, pp. 181–187.

A transistor counting circuit which utilizes large signal techniques and diode switching to achieve versatility and reliability in spite of environmental and supply voltage changes is described. With only a change in polarity, the circuit can utilize either *p-n-p* or *n-p-n* transistors whose dc beta gains vary from 6 to 18.

144
An Emitter-Follower-Coupled, High-Speed Binary Counter, by I. Horn (Burroughs Corp.); 1958 WESCON CONVENTION RECORD, pt. 4, pp. 54–61.

Two versions of a high-speed, nonsaturating binary counter consisting of a nonsaturating, emitter-follower-coupled flip-flop and a complementing network are described. One version uses an unconditional method of complementing and the other uses a conditional method. The former is capable of counting rates of 50 mc when constructed with surface barrier transistors and has the disadvantage of being sensitive to pulse amplitude. The conditional method of complementing eliminates this disadvantage, but prevents using the flip-flop transistors to the limit of their speed capabilities. Both counters are insensitive to changes in pulse width.

145
A Transistor-Magnetic Core Binary Counter, by H. R. Irons (U. S. Naval Ordnance Lab.); PROC. IRE, vol. 46, pp. 1967–1968 (L); December, 1958.

A simple transistor-magnetic core binary counter which is useful at counting rates as high as 2×10^6 pulses per second is described. The counter consists of transistors and cores connected in such a manner that the count is indicated in binary form by the positive and negative residual magnetizations of the cores. The transistors are used as regenerative amplifiers to change the magnetic state of the cores when an input pulse is applied to the counter. A circuit which can provide the proper pulse to operate the first stage of the counter is presented and discussed.

146
Multiplying Circuit Uses Amplifiers, by W. A. Geyger (U. S. Naval Ordnance Lab.); *Electronics*, vol. 32, pp. 58–59; January 9, 1959.

A four-quadrant analog multiplying device in which high-speed magnetic amplifier square-law circuits containing nickel-iron tape cores, silicon diodes, and resistors replace thermal converters is described. Multiplication is based on the relation: $(A+B)^2-(A-B)^2=4AB$. The reversible-polarity output can be used to operate a pen recorder.

147
An Electrical Multiplier Utilizing the Hall Effect in Indium Arsenide, by R. P. Chasmar and E. Cohen (Metropolitan-Vickers Elec. Co., Ltd.); *Electronic Engrg.*, vol. 30, pp. 661–664; November, 1958.

Analog computer multipliers which utilize the Hall effect in the semiconductor indium arsenide are discussed. They consist essentially of a coil wound on a suitable ferromagnetic core with a semiconductor plate mounted in the gap. The advantages of using indium arsenide in this application are pointed out. The design of Hall plates is considered and certain figures of merit are proposed. The construction and performance of a multiplier in which an indium arsenide Hall plate is mounted in the gap in a ferrite core are discussed. The harmonics produced in this multiplier at peak output were less than one per cent of the fundamental.

148
Multiplication by Semiconductors, by C. Hilsum (Services Electronics Res. Lab.); *Electronic Engrg.*, vol. 30, pp. 664–666; November, 1958.

Analog computer multipliers which utilize either the Hall effect or magnetoresistance in semiconductors are described. The possible modes of operation are considered and the choice of materials is discussed. Details of several experimental multipliers and the results obtained from them are given.

149
Diodeless Magnetic Shift Registers Utilizing Transfluxors, by N. S. Prywes (Moore School of E.E., Univ. of Pa.); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 316–324; December, 1958.

The operation of the transfluxor is described briefly and a design procedure for using transfluxors in a two-core-per-bit shift register to provide isolation between stages by magnetic means rather than with diodes is discussed. The transfer of information in such a shift register is analyzed in some detail.

150
Magnetic-Drum Components for High Storage Density, by K. Hong; *Commun. & Electronics*, no. 39 (*AIEE Trans.*, pt. I, vol. 77), pp. 667–672; November, 1958.

In order to augment, correct, and resolve the conflicting viewpoints in the magnetic recording literature, comprehensive formulas for the maximum storage density of a magnetic medium, the playback voltage, and the resolving power of the reading head are derived. Although these formulas are applicable to any magnetic recording, the emphasis is on the magnetic components of a magnetic drum where information is handled in binary digits (bits) and requires a distinction between only two magnetic states, erasing and biasing being unnecessary. It is concluded that the maximum storage density of a medium increases with its coercivity H_c and saturation induction B_s; that in a short-wavelength region the playback voltage is proportional to the H_c of the medium and is independent of its retentivity B_r; and that the resolving power of the reading head is mainly determined by the gap length but is influenced by the field intensity and permeability of the head and the velocity of the medium.

151
A Track Switching System for a Magnetic Drum Memory, by D. D. Majumder (Indian Stat. Inst., Calcutta); *Electronic Engrg.* vol. 30, pp. 702–705; December, 1958.

A system which utilizes a rectifier function mesh and magnetic gating transformers to read from or to write onto a serial type magnetic drum memory of a digital computer is described. The same information channel and magnetic head with one winding can be used for both recording and playing back signals. Many advantages of the system are pointed out.

152
Transistorized Capacitor-Diode Memory System, by H. R. Irons (U. S. Naval Ordnance Lab.); *U. S. Gov. Res. Repts.*, vol. 30, p. 391 (A); November 14, 1958. PB 133950.

A capacitor-diode random access memory system having a capacity of 64 words of 20 bits each is described. The access time is

1 μs and the time required for a complete memory cycle is 7 μs (12 μs when successive memory accesses are to the same work position). The physical size of the unit is comparable to that of a ferrite core-transistor memory, and the power consumption is much less than that of the latter unit.

153
An Analogue Memory, by W. S. Kozak (Canadian Westinghouse Co.); 1958 WESCON Convention Record, pt. 4, pp. 108–122.

The design, frequency response, and applications of an analog computer memory wheel are discussed. In the memory, a high quality capacitor is attached between each of a series of commutator bars arranged along the periphery of a perspex ring and a slip ring. As the wheel rotates at a constant speed, one wiper charges each capacitor to the value of the analog signal and another, at some later time depending on the speed of the wheel and the angular displacement between the two wipers, samples the capacitor voltage levels and feeds the delayed signal back into the computer. With a constant speed of 6 rpm, a maximum delay of 10 seconds is possible. The memory was originally designed for use in a PACE computer rack but is compatible with other types of computers.

154
Manual for Magnetic Tape Input-Output Unit for an ERA 1103 Digital Computer, by P. M. Kintner (Ballistic Res. Labs.); *U. S. Gov. Res. Repts.*, vol. 31, p. 31 (A); January 16, 1959. PB 134284.

A Magnetic Tape Input-Output Unit designed to furnish a means of inserting and withdrawing data into and from an ERA 1103 digital computer is described. The unit will insert data into the computer, utilizing the computer's IOA register and External Read (ER) order, at rates up to 1500 groups per second, one group consisting of eight bits of binary data. Data can be recorded out onto magnetic tape through the unit, again utilizing the computer's IOA register and External Write (EW) order, at rates up to 3000 groups per second.

155
Feasibility and Evaluation Study of a Universal Function Generator, by E. H. Jakubowski (Springfield Armory); *U. S. Gov. Res. Repts.*, vol. 31, p. 22 (A); January 16, 1959. PB 134279.

The design and development of universal function generators for use in analog computing circuits are discussed. In order to simulate operational characteristics of a weapon on an analog computer, several types of function generators are necessary. Circuitry for the elliptical function generator and the ring spring generator is presented.

156
Function Generator for Sines or Cosines, by H. Schmid (Link Aviation, Inc.); *Electronics*, vol. 32, pp. 48–50, January 23, 1959; 1958 WESCON Convention Record, pt. 4, pp. 89–107.

A transistorized trigonometric function generator which provides an output voltage proportional to either the cosine or the sine of the input voltage is described. The generator operates on the principle that the area under the sine curve varies as a cosine function. The input voltage is converted by a linear pulse width modulator into a pulse width which controls an electronic switch. The switch in turn cuts off a portion of the output of a sine-wave source. The output of the switch is averaged by conventional filtering means.

157
Digital Analog Conversion with Cryotrons, by L. K. Wanlass and L. O. Hill (Univ. of California); Proc. IRE, vol. 47, pp. 100–101 (L); January, 1959.

Several methods of using cryotrons as active devices in analog-to-digital and digital-to-analog conversion equipment are described. Two types of cryotron digitizing circuits, cryotron matrices for translating step code to binary code and binary code to step code, and three types of cryotron digital-to-analog converters are discussed. Possible applications of these circuits include their use as input-output devices for large scale digital computers using cryotrons as active elements.

158
BIDEC—A Binary-to-Decimal or Decimal-to-Binary Converter, by J. F. Couleur (G.E. Co.); IRE Trans. on Electronic Computers, vol. EC-7, pp. 313–316; December, 1958.

Simple, high-speed devices to convert binary, binary coded octal, or Gray code numbers to binary coded decimal numbers or vice versa are described. The circuitry required is four shift register stages per decimal digit plus one 30-diode network per decimal digit. In simple form, the conversion requires two operations per binary bit but it is theoretically capable of working at one operation per bit.

159
Digital Recorder for Wind-Tunnel Data, by R. C. MacArthur and W. J. Ungar (Cornell Aeronautical Lab., Inc.); *Electronics*, vol. 31, pp. 86–89; December 5, 1958.

Equipment for recording force and pressure data from wind tunnel models as three decimal digits on punched cards for use with a digital computer is described. This technique provides final data while a tunnel test is in progress. Various basic sections of the equipment can be adapted for digital recording of voltages, currents, and resistances as well as recording frequency and time-coded data.

160
Low Cost Conversion Adapts Univac High-Speed Printer to IBM 704 Outputs, by F. Ketchum (Univ. of California); *Computers and Automation*, vol. 7, pp. 14–16; November, 1958.

A relatively inexpensive method of adapting the Univac high-speed printer to handle the outputs of IBM 704 computers is described.

161
Achieving Maximum Pulse Packing Densities and Transfer Rates, by B. W. Thompson and D. F. Eldridge (Ampex Corp.); 1958 WESCON Convention Record, pt. 4, pp. 48–53.

The problem of increasing the transfer rates of digital computers which utilize magnetic tape is discussed. The various factors which limit the over-all transfer rate in present machines are analyzed and methods for improvement of each of these are suggested.

162
An Input Device for the Harvard Automatic Dictionary, by A. G. Oettinger (Harvard Univ.); *Mech. Translation*, vol. 5, pp. 2–7; July, 1958.

A standard input device which permits transcription of either Roman or Cyrillic characters, or a mixture of both, directly onto magnetic tape is described. The modified unit produces hard copy suitable for proofreading, and records information in a coding system well adapted to processing by a central computer. The coding system and the necessary physical modifications are both described. The design criteria used apply to any automatic information processing system, although specific details are given with reference to the Univac I.

163
A Computer Simulation Chain for Research On Picture Coding, by R. E. Graham and J. L. Kelly, Jr. (Bell Tel. Labs.); 1958 WESCON Convention Record, pt. 4, pp. 41–46.

Equipment for recording a television picture on digital magnetic tape in a suitable form for entry into the IBM 704 digital computer and for playing back a 704 output tape on a television monitor is described. The primary purpose of the whole system is to simulate electronic signal processing or encoding methods proposed for achieving efficient communication. The simulation chain may also be used to analyze pictures for such purposes as pattern recognition or statistical computations. Photographs of pictures showing the system performance are included.

164
Logic for a Digital Servo System, by R. W. Ketchledge (Bell Tel. Labs.); *Bell Sys. Tech. J.*, vol. 38, pp. 1–17; January, 1959.

Methods for performing comparisons of binary numbers are described. These techniques have proved useful in the positioning of crt beams in a photographic memory. A binary address is compared with a digital indication of the present position in circuitry called digital servo logic. The output of the servo logic is an analog indication of the positional error. Logics for obtaining sign only, sign plus magnitude, and sign plus approximate magnitude are described.

165
Catalog of Devices Useful in Automatic Data Reduction, by R. S. Hollitch and A. K. Hawkes (Armour Res. Foundation); *U. S. Gov. Res. Repts.*, vol. 30, p. 576 (A); December 12, 1958. PB 111928r.

The majority of the devices described are digital in nature, and of these, emphasis has been placed on equipment which performs analog-to-digital conversion, that is, voltage-to-digital converters and shaft position encoders, since this class of equipment occu-

pies a leading position in systems for recording data automatically. Also included are digital plotters, printers, digital magnetic tape transports, high capacity memory systems, digital-to-analog converters, airborne magnetic tape recorders, and certain semi-automatic devices, such as oscillogram and film measuring equipment, which produce digital records.

166
Components of Automatic Computing Machinery—List of Types; *Computers and Automation*, vol. 7, pp. 22–24; November, 1958.

Various types of components of automatic machinery for computing or data processing are listed.

A-4: EQUIPMENT—DIGITAL COMPUTERS

167
Automatic Computing Machinery—List of Types; *Computers and Automation*, vol. 7, pp. 20–22; November, 1958.

78 types of machinery that may be considered varieties of automatic computing or data processing equipment are listed.

A-5: EQUIPMENT—ANALOG COMPUTERS

168
General Purpose DC Analog Computer with Transistor Circuitry, by H. L. Ehlers (North American Aviation, Inc.); *Proc. 1958 East Coast Conference on Aeronautical and Navigational Electronics*, pp. 237–248, October 27–28, 1958.

A general purpose analog computer with transistor circuits which is being built for airborne use is described. The system contains stabilized operational amplifiers; resolver, multiplier, and function generator servos; diode function generators; diode multipliers; and other minor auxiliary computing components.

169
Theoretical Consideration of Computing Errors of a Slow Type Electronic Analog Computer, by T. Miura and M. Nagata (Hitachi Central Res. Lab., Tokyo); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 306–312; December, 1958.

Two sources of error inherent in the solution of differential equations by means of analog computers are discussed, namely the integrators and the coefficient setting elements. The error due to the integrators is independent of the order of the differential equation if the integrating time constants are all equal, while the error due to the coefficient setting elements is dependent on the setup procedure. The amount of the error from the coefficient-setting elements is comparatively small in the operating frequency range $\omega < 1$ rad/sec so that in this case, theoretical analysis considering errors due to the integrators only is quite sufficient. However, higher accuracy is required in the range $\omega = 10$ to 100 rad/sec (for example in flight simulators) and consequently the error caused by the coefficient setting elements cannot be neglected. A generalized and practical equation for this error is developed from conventional formulas.

170
Complex Plane Scanner, by C. M. Alaia and P. H. Oden (Columbia Univ.); *U. S. Gov. Res. Repts.*, vol. 30, p. 385 (A); November 14, 1958. PB 134578.

The circuits and the general theory of operation of a specialized type of analog computer, the Complex Plane Scanner, are described. Essentially, this computer will automatically map contours in the complex frequency plane into an arbitrary function plane. A brief résumé of pertinent theory and applications of the computer is presented.

B-1: SYSTEMS—THEORETICAL DESIGN

171
Small Computers for Inertial Navigation, by H. R. Brown (North American Aviation, Inc.); *Proc. 1958 East Coast Conference on Aeronautical and Navigational Electronics*, pp. 129–135; October 27–28, 1958.

The use of small transistorized analog and digital computers as components in inertial navigation systems is discussed. Analog computers are used in the shorter range and special-purpose applications while digital differential analyzers (DDA) and combination DDA-General Purpose (GP) computers are used in the longer range and multipurpose applications. Simplification and size-reduction of the vehicular-mounted equipment are primary requirements and are illustrated by examples of such equipment.

172
Effect of Quantization in Sampled-Feedback Systems, by J. E. Bertram (Columbia Univ.); *U. S. Gov. Res. Repts.*, vol. 31, p. 19 (A); January 16, 1959. PB 133341.

The rapid advance of digital computers and digital technology has in recent years resulted in the development of digital versions of practically every control system component. As a result, mixed systems composed of both digital and analog elements are becoming quite common. The effect of quantization or round-off errors, inherent in the digital elements, on the performance of feedback control systems is discussed.

173
Automatic Techniques, Large Computers, and Engineering Calculations, by V. Paschkis (Columbia Univ.); IRE TRANS. ON INDUSTRIAL ELECTRONICS, vol. PGIE-7, pp. 27–32; August, 1958.

The three classes of large-scale computing devices—the digital computer, the differential analyzer, and the analog computer—are briefly described and their application to automation is discussed. In preparing for automation, the computers can be used to analyze a manufacturing process thoroughly. A result of this analysis is a less complex automation computer (the computer which automatically sets the process controls in the automatized factory). Several examples of engineering calculations which the computer can perform are given. The organization of a computing laboratory for engineering calculations is discussed and the social implications of automation are considered.

174
Dynamic Systems Synthesizer, by E. C. Hutter, J. Lehmann, and others (RCA); *U. S. Gov. Res. Repts.*, vol. 30, p. 386 (A); November 14, 1958. PB 151137.

A study of the requirements for a new modern analog computer facility that would be sufficient to simulate modern guided missile systems is described. A number of new components were developed to enable the proposed computer to operate on a 1:1 time scale. These new components and the proposed new programming system have been tested on a model computer that contains one or more of all the components. These tests indicate that the methods and components suggested are sound, and that their use would result in a guided missile simulator that could operate in real time.

175
The Design of Analog Computer Compensated Control Systems, by S. C. Bigelow (Columbia Univ.); *Applications and Industry*, no. 39 (*Trans. AIEE*, pt. II, vol. 77), pp. 409–415; November, 1958.

A procedure for the design of a tandem compensated duplicator control system which utilizes an analog computer to realize the transfer function of the compensating system is presented. Restrictions on cancellation of poles and zeros of the plant by zeros and poles of the compensating system and explicit relations between steady-state system error or error constants and the coefficients of the over-all system transmission function which hold for any design system are obtained.

176
Fundamental Concepts in the Theory of Systems, by B. P. Sauer (Univ. of Chicago); *U. S. Gov. Res. Repts.*, vol. 31, p. 30 (A), January 6, 1959. PB 151242.

An extensive report on some fundamental concepts in the theory of systems is presented. Section I contains an expository discussion of digital processes including a brief historical account of the conception of routine computational procedures in classical mathematics, an exposition of the Post-Turing analysis of uniform digital computation and of the so-called Universal Turing Machine, and the abstract idea of a digital process. It concludes with an example of self-reproducing behavior by a digitally organized process. Section II, "Discrete Linear Mechanisms," and Section III, "Continuous Mechanisms," treat several topics in the basic theory of pulsed and continuous servomechanisms. Section IV contains an extensive development of weighing functions from the point of view of functional analysis. In Section V, an important class of mechanisms known as finite automata is analyzed with the help of algebraic techniques. The main theorem (due to Kleene) provides an almost complete characterization of the tasks such mechanisms can perform. A number of formulations of this characterization are proved equivalent.

B-2: SYSTEMS—DESCRIPTION

177
Transac C-1100: Transistorized Computers for Airborne and Mobile Systems, by G. L.

Hollander (Philco Corp.); IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-5, pp. 159–169; September, 1958.

The Transac C-1100 digital control computers designed to handle in a single computer such functions as aircraft stabilization, navigation, cruise control, and landing are discussed. A functional description of the computers is presented and their program structure is discussed. The general-purpose stored-program organization allows the C-1100 to be shared by various tasks during flight, and enables the same computer to execute a large variety of operations merely by changing the program. A powerful system of internal and external decisions permits the computer to respond to both the en route instructions by the pilot and to conditions of the system as sensed in one of the 64 data channels. A problem involving dead-reckoning and VORTAC navigation, automatic control of the aircraft, and fuel management is given to illustrate the computer programming. The physical layout and packaging of a typical computer are described and some auxiliary equipment is briefly discussed. A typical transistorized C-1100 computer has a precision of one part in 10^6 and can perform additions at the rate of 60,000 per second.

B-4: SYSTEMS—TESTING

178
A Library of Blip Samples for Use in the Realistic Simulation and Evaluation of Automatic Radar Data Processing Systems, by C. M. Walter and H. M. Willett (Air Force Cambridge Res. Center); 1958 WESCON CONVENTION RECORD, pt. 4, pp. 8–27.

The use of a combined special purpose analog and general purpose digital computer configuration, operating in real time, to obtain a library of radar video blip samples is discussed. The library consists of a number of blip samples on punched paper tape. Each sample consists of a long sequence of blips, drawn from a stationary population, together with the more significant statistical parameters characterizing the sample. Samples exist for a number of radar systems parameters such as antenna gain pattern, axial signal-to-receiver noise ratio, target scintillation spectra, and pulses-per-beam width. The availability of the library is cited and several applications for it are discussed in the evaluation.

D-1: PROGRAMS—AUTOMATIC PROGRAMMING, DIGITAL COMPUTERS

179
Preliminary Report—International Algebraic Language, by the ACM-GAMM Committee on Algebraic Programming Language; *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 8–22; December, 1958.

A preliminary report on a proposed international algebraic programming language is presented. The stated objectives of the outlined reference language are: 1) it should be as close as possible to standard mathematical notation, readable with little further explanation; 2) it should be usable for describing computer processes in publications; 3) it should be mechanically translatable into machine programs. The proposed algorithmic language is developed in great detail, the main division being into basic symbols, expressions, statements, and declarations.

180
The Problem of Programming Communication with Changing Machines, Part 2, by the Share Ad Hoc Committee on Universal Languages; *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 9–15; September, 1958.

This continues, with three detailed examples of the use of UNCOL, the development of the three-level approach to the problem outlined in pt. 1 [abstract 63]. The first level consists of all POL, or problem-oriented languages, translatable by generating routines into UNCOL, or universal computer-oriented language. Further translation routines go from UNCOL to any specified machine language, ML. A concise system of sub- and super-scripts indicates which language any routine is written in, and which two languages it connects. The scheme, if fully developed, could translate a program written for any specific machine into a program written for any comparable machine, without the interaction of any human programming.

181
Generalization: Key to Successful Electronic Data Processing, by W.G. McGee (G.E. Co., Hanford Atomic Products Operation); *J. Assoc. for Computing Mach.*, vol. 6, pp. 1–23; January, 1959.

The need for the generalization of the commoner procedures encountered in data-processing, with the consequent savings in programming manpower, is stressed. Three such routines: generalized sorting, file maintenance, and report generation (or extraction of information from files) are described. In addition to reducing cost, these routines permit the mechanization of comprehensive source files and provide unprecedented flexibility in the design of new data processing systems.

182
On Matrix Program Schemes, by Iu. I. Ianov; *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 3–6; December, 1958 [Translated from *Dokl. AN USSR*, vol. 113, pp. 283–286; 1957].

An abstract matrix theory of operations is developed. A finite set of operations $A_1, \cdots, A_n$ has its order of performance specified by the values of $(n+1)$ functions of k two-place logical variables $p_1, \cdots, p_k$. All the orders of performing the operations can be written as a matrix whose elements are specific logical functions of the p_i. These matrices are studied with an eye to their application in the theory of programming schemes.

183
A Programming Language for Mechanical Translation, by V. H. Yngve (Mass. Inst. Tech.); *Mech. Translation*, vol. 5, pp. 25–41; July, 1958.

A notational system for use in writing translation routines and related programs is described. The system is specially designed to be convenient for the linguist so that he can do his own programming. Programs in this notation can be converted into computer programs automatically by the computer. Complete instructions for using the notation are presented and some illustrative programs are shown.

184
The Pegasus Autocode, by B. Clarke and G. E. Felton (Ferranti Ltd.); *Computer J.*, vol. 1, pp. 192–195; January, 1959.

This autocode is one of a growing family of compiling and pseudo-codes, whose object is to simplify the programming process for someone not familiar with the detailed operation of a particular machine. Most of the common mathematical operations are covered by one word of the pseudo-code. The scheme is particularly useful for "one-off" programs (*i.e.*, programs that are used only once), and as a starting-off point in the preparation of more detailed programs.

185
"DEUCE" Interpretative Programs, by C. Robinson (Eng. Elec. Co. Ltd.); *Computer J.*, vol. 1, pp. 172–176; January, 1959.

Three interpretative programs for use with the DEUCE Computer are described. These are used for the coupling, editing, and simplification of programs and bear marked similarities to such better-known schemes as FORTRAN. The coder breaks down his problem into a series of "three-address, one function" code words. The system is particularly suited to problems where time of operation can take second place to ease of programming and testing; although where large blocks of data are concerned, time of operation does not suffer unduly.

D-2: PROGRAMS—APPLICATIONS, DIGITAL COMPUTERS

186
Algorithms and the Machine Decision Problem, by B. A. Trachtenbrot; *U. S. Gov. Res. Repts.*, vol. 30, p. 384 (A); November 14, 1958 [Translated from *Matematika b Shkole*, vol. 3, no. 4, pp. 3–10, and no. 5, pp. 5–14]. PB 132489.

A report on the development of electronic digital computers and their use for solving mathematical and logical problems is presented. The sphere of use of these computers grows continually. They solve complicated mathematical problems requiring study and solution of very cumbersome systems of algebraic or differential equations; they translate text from one language to another, play chess, etc. There exists the prospect of using them as production devices to guide automation processes.

187
Computers and Commerce: 4—Management and Control, by A. S. Douglas (Univ. of Leeds); *Computer J.*, vol. 1, pp. 168–171; January, 1959.

It is shown by means of examples how the method of Operations Research known as Linear Programming may be used to provide mathematical models, tractable on computers, of situations of interest to business management, such as the control of distribution and production. A plea to British management to take a more enterprising view of the tools being developed by computer technology is included.

188
Automation and the Office, by H. W. Gearing (Metal Box Co., Ltd.); *Computer Bull.*, vol. 2, pp. 43–46, October/November, 1958; pp. 59–63, December, 1958/January, 1959.

Several reasons for the slow arrival of full automation in the office are discussed.

189
Computers for the Highway Engineering Program, by J. Belzer (Battelle Memorial Inst.); *Computers and Automation*, vol. 8, pp. 12–13; January, 1959.

The use of electronic computers to relieve engineers of tedious and costly computations in the design of highways and bridges is discussed.

190
Distribution System Primary-Feeder Voltage Control, by H. E. Lokay, D. N. Reps, and G. J. Kirk, Jr. (Westinghouse Elec. Corp.) and H. K. Amchin and R. J. Bentzel (Amer. Elec. Power Service Corp.); *Power Apparatus and Systems* (*AIEE Trans.*, pt. III, vol. 77), pp. 845–879; October, 1958.

The use of a digital computer to determine the optimum combination of voltage regulators and capacitors for primary feeder voltage controls in an electrical distribution system is described. The computer is programmed to decide automatically which of several alternative combinations of voltage control equipment are applicable for a given primary-feeder circuit based on stipulated line-voltage regulation limits; to determine the required equipment locations and ratings; and to compute the cost of each feasible alternative, including such economic factors as equipment installed cost, system investment, and I^2R losses. A step-by-step description of the procedure is given, the analytical preparation and contents of the digital computer program is discussed, and the practical application of the method to actual systems is described.

191
Digital Computer Aids in Power-Pool Operation Studies, by H. M. McIntyre, C. W. Blake, and J. S. Clubb (Bonneville Power Administration); *Commun. & Electronics*, (*AIEE Trans.*, pt. I, vol. 77), pp. 652–657; November, 1958.

The use of a general-purpose digital computer as an aid in the solution of problems encountered in the pooled operation of a large hydroelectric system is described. The program developed for this purpose has enabled studies of operations to be prepared more rapidly and in a form that can be used readily by management in analyzing current conditions and in formulating plans for subsequent operations.

192
Solving Equations of State in Telephone Traffic Theory with Digital Computers, by S. G. Carlsson and A. Elldin (Telefonaktiebolaget LM Ericsson); *Ericsson Technics*, vol. 14, no. 2, pp. 221–224; 1958.

The problem of evaluating large systems of equations of state in telephone traffic theory is discussed, and a linear iteration method which has been applied to the equations of state for a three-group grading with random hunting is described. Evaluations carried out on the Swedish digital computer BESK using this method have given very promising results. An equation system with 729 unknowns was solved in the machine to 6 decimal places (average 4 correct digits) in 7 minutes. The iteration method, which is applicable to most types of ordinary equations of state in statistical equilibrium, can probably be used to solve systems with at least 1000 or 2000 unknowns within reasonable evaluation times.

193
Artificial Traffic Trials on a Two-Stage Link System Using a Digital Computer, by B. Wallstrom (Telefonaktiebolaget LM Ericsson); *Ericsson Technics*, vol. 14, no. 2, pp. 259–289; 1958.

A series of Monte Carlo experiments designed to study the traffic capacity of a two-stage link system under various conditions is described. These experiments have been performed using the Swedish digital computer BESK. An approximate formula for the accuracy of the time congestion estimates is derived, and an exact calculation of congestion for a symmetrically loaded link system with only one outlet per route is carried out. The following aspects have been studied in the traffic trials: 1) the influence of various numbers of inlet columns and routes; 2) the influence of the hunting rules (random and sequential hunting); and 3) the accuracy of some known calculation methods.

194
Technical Studies in Cargo Handling, Part I: Formulation of Recurrence Equations for Shuttle Process and Assembly Line, by R. Bellman (Univ. of California); *U. S. Gov. Res. Repts.*, vol. 30, p. 461 (A); November 14, 1958. PB 129917.

The recurrence technique is applied to two mathematical models of cargo handling, the "link-node" model and the assembly line model. The object is to derive the basic equations which may be used to describe the process.

195
Technical Studies in Cargo Handling, Part II: Computation of Delays in the Multistage Shuttle Process, by R. Bellman, Y. Fukuda, and M. Pollack (Univ. of California); *U. S. Gov. Res. Repts.*, vol. 30, p. 461 (A); November 14, 1958. PB 133217.

A Monte Carlo approach to the calculation of delays in the multistage shuttle process by means of SWAC, a high-speed digital computer, is described. Several codes were developed for SWAC to generate the random time elements and to calculate the delays in the 2nd stage for 3-, 4-, 5-, and 6-stage shuttle processes. It was found that the 2nd stage delays did not seem to be influenced by the item number but were affected slightly by the number of stages, the delays tending to increase with increasing number of stages.

196
Technical Studies in Cargo Handling, Part III: Distribution of Delay in the Three Stage Shuttle Process, by Y. Fukuda (Univ. of California); *U. S. Gov. Res. Repts.*, vol. 30, p. 461 (A); November 14, 1958. PB 132627.

An attempt to analyze the delays found in the three stage shuttle process is described. Owing to the special character of these recurrence relations between the delays, it is feasible to derive the exact distribution of every delay which occurs in transportation of each unit commodity. Several numerical examples are given, and the results are illustrated by means of graphs together with the results of SWAC computation.

197
Technical Studies in Cargo Handling, Part IV: Methods of Computing Delays in an *N*-Stage Shuttle Process, by M. Pollack (Univ. of California); *U. S. Gov. Res. Repts.*, vol. 30, p. 461 (A); November 14, 1958. PB 133230.

Two basic approaches which may be used as a form for the computation of delays in the N-stage shuttle process with random working times are discussed. The two simulation and recurrence equations are each investigated for use on a high-speed digital computer.

198
Aircraft Route Analysis on a Digital Computer, by N. G. Moorhead (formerly of Vickers-Armstrongs (Aircraft) Ltd.); *Computer J.*, vol. 1, pp. 160–162; January, 1959.

A computer program for examining the performance of civil aircraft over specific routes with a view to optimizing the routes is described. The program has the desirable features that all special requirements (tail winds, initial and final cruises, etc.) are anticipated, and that the final analysis is printed out in a form suitable for use without manual typing. Further sophistication of the program is feasible should the need arise.

199
Applying Computers to Air Traffic Control, by H. S. Stokes (Airways Modernization Board); IRE Trans. on Aeronautical and Navigational Electronics, vol. ANE-5, pp. 152–159; September, 1958.

Following a brief description of the present manual system of air traffic control and an outline of the Curtis program for modernization of our national airways system, the plans of the Airways Modernization Board for applying computers to air traffic control are discussed. Both en route and terminal area control systems are considered.

200
Solution of Algebraic and Transcendental Equations on an Automatic Digital Computer, by G. N. Lance (Univ. of Southampton); *J. Assoc. for Computing Mach.*, vol. 6, pp. 97–101; January, 1959.

Three methods of solving the equation

$$f(z) \equiv f(x + iy) \equiv \phi(x, y) + i\psi(x, y) = 0$$

by evaluating the minima of the function $S \equiv |\phi| + |\psi|$ are described. Method I (alternate adjustment of variables) starts with the selection of an arbitrary point $P(x_0, y_0)$ and attempts to reduce S by alternately adjusting the values of x and y. This method is inadequate as it does not always lead to convergence. Method II is the well-known one of steepest descent. Each new step of the solution is taken in the direction of the gradient vector

$$\left(-\frac{\partial S}{\partial x}, -\frac{\partial S}{\partial y}\right).$$

Method II always converges. However, not only S, but its partial derivatives, must be calculated at each step. Method III removes this difficulty by proposing difference equation approximations to the partial derivatives.

201
Iterative Predictor Selection for a Single Criterion, IBM 650 Computational Program, by C. E. Lunneborg (Washington Univ.); *U. S. Gov. Res. Repts.*, vol. 30, p. 388 (A); November 14, 1958. PB 132647.

A computer program which will process matrices of predictor intercorrelations and vectors of validity coefficients of order as high as 45 is described. The computations carried out by the program are those described by Horst with one exception. The Wherry shrinkage formula for multiple correlation coefficients has been replaced in this program by a shrinkage formula suggested by Snedecor.

202
Stable Predictor-Corrector Methods for Ordinary Differential Equations, by R. W. Hamming (Bell Tel. Labs.; *J. Assoc. for Computing Mach.*, vol. 6, pp. 37–47; January, 1959.

A general technique of developing generalized predictor and corrector formulas based on undetermined coefficients is described. The characteristic roots of the difference equation give criteria for stability. The main cost of gaining stability is the loss of some accuracy. In the principal case treated, this can be compensated by shortening the interval of integration by about 15 per cent.

203
Design of Multilayer Filters by Successive Approximations, by P. Baumeister (Univ. of California); *J. Opt. Soc. Amer.*, vol. 48, pp. 955–958; December, 1958.

The application of an IBM 650 MDDPM digital computer to the design and fabrication of optical filters is discussed. A relaxation method is used to adjust the thickness of the films of a multilayer in order to alter its transmission characteristics in a limited spectral region. This method is used to modify the spectral transmission of a short-wavelength pass multilayer and a broad-band dielectric mirror. Since the computer indicates the effect of the thickness of each layer on the transmission of a multilayer of various wavelengths, an operator can readily determine whether or not to continue the fabrication process when an error is made in the thickness of an evaporated layer. If he should decide to continue, this same information can permit him to alter the thickness of the subsequent layers in order to correct the error.

204
Dyes—Mixture Analysis with an Electronic Calculating Punch, by J. L. F. deKerf (Gevaert Photo-Production N.V.); *J. Opt. Soc. Amer.*, vol. 48, pp. 972–975; December, 1958.

The calculation of the tristimulus specification of subtractive dye mixtures consists of three parts: the calculation of the spectral density of the mixture, the conversion to the spectral reflectance or transmittance, and the deduction of the tristimulus values and chromaticity coordinates. With a standard IBM 604 electronic calculating punch, three runs are needed and the yield is about 24 to 60 color points an hour, according to the desired accuracy. It is shown that with an expanded type, the 604-004, the complete calculation can be performed in a single run. The yield attained in this way is 60 to 150 color points an hour.

205
Automatic Data Reduction of Spot Diagram Information, by W. E. Goetz and N. J. Woodland (IBM Corp); *J. Opt. Soc. Amer.*, vol. 48, pp. 965–968; December, 1958.

A technique whereby modern electronic computers can produce spot diagram data at at pace which makes human selection of best spot diagrams impractical is described. In a particular mechanical image dissector, the energy falling onto the photocathode redistributes as the scan progresses from field point to field point. The reviewing surface is considered to be a rectangular network. Rays representing unit energy are traced through the system, and the total energy in each square is determined for each field point. Applying chi-squared methods to the numerical matrices derived for the field points in a scan, the redistribution between field points is evaluated numerically. This digital evaluation makes feasible automatic data reduction of spot diagram information.

206
Crystal Structure Refinement by Least Squares with the IBM 650, by L. R. Lavine and J. R. Steinberg. (Mass. Inst. Tech.); *U. S. Gov. Res. Repts.*, vol. 30, p. 394 (A); November 14, 1958. PB 132531.

A program for performing least-squares refinements of crystal structure is described. Details of subroutines of possible interest to others are included.

207
An Application of Electronic Computing to X-Ray Crystallography, by J. Gillis (Weizmann Inst. Sci., Israel); *Acta Crystallographica*, vol. 11, pp. 833–834; December, 1958.

A program for examining observed structure factors for significant Harker-Kasper inequalities which has been run successfully on the WEIZAC electronic digital computer is described. As an incidental aid in the preparation of the program, a polynomial approximation to an atomic scattering factor of carbon has been calculated.

208
Application of the Monte Carlo Method to the Lattice-Gas Model I. Two-Dimensional Triangular Lattice, by Z. W. Salsburg (Rice Inst.) and J. D. Jacobson, W. Fickett, and W. W. Wood (Univ. of California); *J. Chem. Phys.*, vol. 30, pp. 65–72; January, 1959.

Application of the Monte Carlo numerical method for obtaining statistical mechanical averages in the petite canonical ensemble to the two-dimensional triangular lattice-gas model is described. The program has been carried out on an IBM 704 and has been limited to two-dimensional lattices with nearest-neighbor interactions.

209
Solution of Compressible Boundary Layer Problems by a Finite Difference Method, Part II: Further Discussion of the Method and Computation of Examples, by D. C. Baxter and I. Flugge-Lotz (Stanford Univ.); *U. S. Gov. Res. Repts.*, vol. 30, p. 427 (A); November 14, 1958. PB 133489.

The solution of compressible boundary layer problems by a finite difference method is discussed, and the results of some sixty examples which have been computed using the difference method and a digital computer are given. These results provide some insight into compressible flows with variable pressure gradients and surface temperatures. They illustrate certain trends which can be expected and allow comparison with various approximate methods of solution which have been proposed in the past.

210
Evaluation of the Umkehr Effect by Means of a Digital Electronic Computer, by H. U. Dütsch (Lichtklimatisches Observatorium, Switzerland); *U. S. Gov. Res. Repts.*, vol. 30, p. 577 (A); December 12, 1958. PB 134507.

The nonlinear problem of computing the distribution of the atmospheric ozone into nine layers has been linearized with the help of tables which give the influence of ozone changes in a single layer on the Umkehr curve. These tables and a theoretical Umkehr curve belonging to a standard ozone distribution were calculated by an electronic computer.

211
Monte Carlo Method for Computing the Transmission of Fast Neutrons through a Lead Shield, by J. T. Humphries (U. S. Air Univ.); *U. S. Gov. Res. Repts.*, vol. 30, p. 603 (A); December 12, 1958. PB 134685.

A digital computer program employing Monte Carlo techniques which was designed to compute the transmission of fast neutrons through a spherical lead shield is described. The source is contained in a void at the center of the shield and surface fluxes are computed. Both elastic and inelastic scattering of the neutrons are considered by the program. Flow charts have been constructed to facilitate coding.

212
Abstracts—Nuclear Reactor Codes, by V. Nather and W. Wangren (General Atomics); *Commun. Assoc. for Computing Mach.*, vol. 2, pp. 6–32; January, 1959.

Abstracts of 238 programs (codes) used in the design of nuclear reactors are presented. The programs are classified into eight groups with the following titles: burn-up, engineering, group diffusion, kinetics, miscellaneous, Monte Carlo, physics, and transport. Each abstract has the following form:

Program name (category—installation—machine)

1) Author(s)/person to contact
2) Present status
3) Statement of problem and program description
4) Estimated computing times
5) References
6) Limitations and special comments.

213
Prediction of the Thermal Behavior of Blast-Cooled Generators by Use of a Digital Computer, by R. M. Moroney (Mass. Inst. Tech.); *Applications and Industry*, no. 39 (*AIEE Trans.*, pt. II, vol. 77), pp. 388–394, November, 1958; *U. S. Gov. Res. Repts.*, vol. 30, p. 375 (A); November 14, 1958. PB 134639.

The programming of a digital computer to compute the electrical rating of any blast-cooled aircraft generator, given a complete description of the machine geometry, is discussed. Such a computer routine can be used to predict rating data for machines in the design stage. Promising agreement between computed and measured data has been obtained for two machines. For complete generality of the computer routine, however, heat-transfer equations more accurate than those presently available are required.

214
Computer Analysis of AC Aircraft Generators, by J. R. M. Alger, E. F. Magnusson, and J. T. Duane (G.E. Co.) and R. T. Smith (Univ. of Texas); *Applications and Industry*, no. 39 (*AIEE Trans.*, pt. II, vol. 77), pp. 394–399; November, 1958.

The use of a stored program digital computer in the design of ac aircraft generators is described. The performance analysis program developed for this purpose quickly and accurately provides complete data on conventional generators as dictated by system requirements. Unusual aspects of program use and flexibility, considering the output data desired and the special forms of data presentation required, are discussed. Other topics considered include the machine theory involved, the use of computer logic circuitry, and the general results obtained.

215
Digital Calculation of Network Functions Used in Loss Formula Studies, by R. W. Ferguson, R. W. Long, and L. J. Rindt (Westinghouse Elec. Corp.); *Commun. & Electronics*, no. 39 (*AIEE Trans.*, pt. I, vol. 77), pp. 647–652; November, 1958.

Digital computer programs for calculating the self and mutual drops of a system and for combining the drops of two or more systems to form the drops for the system composed of the combination of these subsystems are described. A ramification of the second program is the inclusion of a new line or group of lines into the matrix of self and mutual drops for a basic system. The digital computer method is more accurate and more rapid than the ac network calculator method.

216
Predistorted Filter Design with a Digital Computer, by P. R. Geffe (Audio Dev. Co.); 1958 WESCON Convention Record, pt. 2, pp. 10–22.

An IBM 650 digital computer program for the design of Butterworth and Tchebycheff filters and for some more general applications is described. Within limits the program will predistort and design any all-pole network for which the complex transfer function is known. In the case of Butterworth and Tchebycheff filters, it will also calculate the transfer function from the desired pass-band ripple factor and the order of the network.

217
Simulation of a Human Tracking Problem of the UDEC III Computer, by H. L. Platzer (Burroughs Res. Center); 1958 WESCON Convention Record, pt. 4, pp. 286–291.

The simulation on a UDEC III digital computer of a man-machine tracking system containing analog and digital components is discussed. The simulation was carried out to obtain word-length requirements for the digital portion of the system, and to study the performance and stability of the complete system. Since the simulation study used digital techniques, simulated the man, and used a unique method for solving the differential equations of the system, it could be performed at an advance time scale on a high speed digital computer. The results of the simulation are given.

218
Team Approach to Computer Programs for Numerical Control, by E. F. Carlberg (Boeing Airplane Co.); *Control Engrg.*, vol. 6, pp. 77–80; January, 1959.

It is pointed out that there are several possible approaches to the problem of developing a computer program for numerical control. For example, the program may be predominantly oriented toward the computer, the machine tool to be controlled, the mathematics of the problem, the drawings of the parts to be machined, or the language involved in communicating with the computer. A team approach to the development of a universal program which avoids the limitations of these individual approaches is discussed.

219
The Use of Computers in Inspection Procedures, by E. Muller (Princeton Univ. and IBM Corp.); *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 7–18; November, 1958.

A general programming procedure for selecting samples from a population classified into subsets is outlined. The problem is to select a stratified random sample (*i.e.*, a specified number of terms at random from each set). Provision is made for the case where certain specified items need to be inspected 100 per cent. A flow-diagram is given indicating inputs, editing and updating, generation of random numbers (by "Residue Classes"), selection of requisite sampling plan, and output. The program may be augmented for feedback evaluation of the inspection technique used.

220
Data Preparation for Numerical Control of Machine Tools, by H. D. Huskey and D. E. Trumbo (Univ. of California and Bendix Aviation Corp.); 1958 WESCON Convention Record, pt. 4, pp. 3–7.

The preparation by means of a computer of the punched tape input which is used to control a precision three axis mill capable of pocket or contour milling is discussed. A part whose boundary is described by any second degree equation can be cut. If the boundary of the part is not a simple one, the computer solves five linear equations in five unknowns to determine the coefficients of the second degree equation. From a given starting point another point on the boundary a short distance away is found by a Taylor series expansion in two variables. From this point, the required offset for the cutter center path is found. These points are spaced close enough together so that a straight line cut between them gives a sufficiently good approximation to the true curve.

221
Chemical Structure Searching with Automatic Computers (Natl. Bur. of Standards); *Computers and Automation*, vol. 7, pp. 16–17; November, 1958.

A program to develop automatic techniques of information storage and retrieval and to apply them to problems of patent search is discussed briefly, and an experimental application of such techniques to generic searches through files of chemical structure diagrams is described. The experiments indicate that an automatic data processing system can be used successfully for very rapid scanning of a file and also in various auxiliary operations.

222
Research Methodology for Machine Translation, by H. P. Edmundson and D. G. Hays (The RAND Corp.); *Mech. Translation*, vol. 5, pp. 8–15; July, 1958.

The methods now in use at the RAND Corporation for research on machine translation of scientific Russian are described. The general approach is that of convergence by successive refinements. The philosophy that underlines this approach is empirical. Statistical data are collected from careful translation of actual Russian text, are analyzed, and then are used to improve the program. Text preparation, glossary development, translation, and analysis are described.

223
The Use of Punctuation Patterns in Machine Translation, by G. Salton (Harvard Univ.); *Mech. Translation*, vol. 5, pp. 16–24; July, 1958.

The analysis of English language punctuation patterns is discussed and a set of specifications for an automatic punctuation analysis program is presented. It is believed that these specifications are capable of treating adequately about 95 per cent of the word strings likely to be encountered in technical texts and that they contain the main bulk of the results which can be achieved by punctuation analysis. A text is analyzed to show the possibilities inherent in the proposed method.

224
A Method for Synthesizing the Waveform Generated by a Character, Printed in Magnetic Ink, in Passing Beneath a Magnetic Reading Head, by I. Flores and F. Ragonese (Remington Rand UNIVAC); IRE Trans. on Electronic Computers, vol. EC-7, pp. 277–282; December, 1958.

A method for determining the waveform generated when a character printed in magnetic ink passes beneath a magnetic reading

head from the geometry of the printed character is described. The character shape is divided into elementary vertical units and the height of these units is then tabulated. A single set of experimental data is obtained in the laboratory by passing a magnetically printed bar beneath the reading head which is used to read the characters. Formulas for combining the geometrical data obtained from the character with the experimental data obtained from the laboratory run with the head are derived. A method for programming a high-speed digital computer to derive the waveforms is then described. These waveforms are often superior to those obtainable in an actual laboratory run of the printed characters.

225
Description and Results of Experiments With Speech Using Digital Computer Simulation, by E. E. David, Jr., M. V. Mathews, and H. S. McDonald (Bell Tel. Labs.); 1958 WESCON CONVENTION RECORD, pt. 7, pp. 3–10.

The digital simulation of operating models of speech coding and transmission schemes on general purpose computers is discussed. The input to the computer is prepared by a converter which samples, quantizes, and records actual speech on IBM 704 magnetic tapes. The converter also reproduces speech from tapes prepared by the computer. Computer requirements and the characteristics of programming peculiar to this type of simulation are discussed. A coding scheme which was examined in detail by this method is described. The required channel capacity and memory size, and the intelligibility of the output speech were determined.

D-3: TECHNIQUES, DIGITAL COMPUTERS

226
DEPI: Interpretive Digital-Computer Routine Simulating Differential-Analyzer Operations, by F. H. Lesh and F. G. Curl (Calif. Inst. Tech.); *U. S. Gov. Res. Repts.*, vol. 31, p. 31 (A); January 16, 1959. PB 133820.

A routine simulating the operation of an analog computer, constructed for use on the digital computer at the Jet Propulsion Laboratory, is described. This code, called DEPI as an abbreviation for differential equations pseudo-code interpreter, provides much of the flexibility and ease of programming associated with analog-computer operation. A study in which the speed, accuracy, and flexibility of the DEPI system were compared with these attributes in other computers and computer programs is reported. The vehicle of comparison was a standard missile-trajectory problem.

D-4: TESTING, DIGITAL COMPUTERS

227
Test Routines Based on Symbolic Logical Statements, by R. D. Eldred (Datamatic); *J. Assoc. for Computing Mach.*, vol. 6, pp. 33–36; January, 1959.

A successful test routine to guarantee that a system has no faulty components should be devised at the level of the components rather than at the level of the programmed orders. A logical method for selecting the minimum number of test conditions necessary to check all the operations of components in a gating structure is described. If a function contains more than one gate, the tests are applied to each gate individually. In an application on the Datamatic 1000 Central Processor, the method supplied a program of approximately 110 test orders that checked 900 tubes and 21,000 diodes for all conditions of activation and inhibition in about 0.08 second.

228
Automatic Failure Recovery in a Digital Data Processing System, by R. H. Doyle, R. A. Meyer, and R. P. Pedowitz (IBM Corp.); *IBM J. Res. & Dev.*, vol. 3, pp. 2–12; January, 1959.

A program which enables a complex digital data processing system to give "first aid" to itself is described. The FIX program automatically compensates for computer malfunctions so that recovery from errors may be effected with a negligible loss of operational time. Some equipment features used by the FIX program are briefly outlined and the structure and function of the program itself are described in detail. In its initial application in the SAGE system, FIX provided automatic recovery from more than 90 per cent of all failures occurring during the period studied.

D-5: PROGRAMS—APPLICATIONS, ANALOG COMPUTERS

229
Matrix Programming of Electronic Analog Computers, by R. E. Horn (Westinghouse Elec. Corp.) and P. M. Honnell (Washington Univ.); *Commun. & Electronics*, no. 39, (*AIEE Trans.*, pt. I, vol. 77), pp. 420–428; September, 1958.

The advantages of matrix programming of electronic analog computers are indicated by means of specific practical examples. It is shown that matrix programming of differential equations organizes the equations, thereby minimizing chances for errors; that it simplifies the scaling of coefficients and variables to suit the characteristics of the available machine; that in many instances, it permits the reduction of the number of "sign changers"; and that it helps safeguard against the appearance of extraneous terms in the solutions of certain systems of differential equations.

230
Study of the Resonating Yawing Motion of Asymmetrical Missiles by Means of Analog Computer Simulation, by J. A. M. Schmidt (U. S. Ballistic Res. Labs.); *U. S. Gov. Res. Repts.*, vol. 30, p. 597 (A); December 12, 1958. PB 134637.

A study made with the aid of an analog computer of the yawing motion of asymmetrical missiles whose rolling velocities are in or near the resonance region is described. The study was undertaken to determine whether this motion is adequately described by the linearized theory. The technique of curve fitting is described and qualitative and quantitative results are presented.

231
Diffusion Concentration Profiles by Analog Computation, by W. Waring (Raytheon Manufacturing Co.); *J. Electrochem. Soc.*, vol. 105, pp. 695–699; December, 1958.

Simulation of diffusion processes on an analog computer for boundary conditions which make analytical solutions impractical is described. Effects on transistor structure are shown for "slumping" and "outdiffusion" of the first diffusant in a double diffusion. Diffusion from a liquid alloy dot through a regrown region is well represented by an analytical solution except for very thin regrown regions or very long diffusion times. Diffusion through two layers of material with different diffusion coefficients gives profiles which may deviate widely from the error function complement.

232
Steam-Generator Instrumentation for a Nuclear Power Plant with Analog Computer Verification of Dynamic Action, by E. E. Lynch and D. P. Waite (G.E. Co.); *Commun. and Electronics*, no. 39 (*AIEE Trans.*, pt. I, vol. 77), pp. 690–696; November, 1958.

The use of an analog computer to study oscillation and dynamic characteristics of steam-generator instrumentation systems for a nuclear power plant is described. The computer setup eventually includes newly developed physical equipment which replaces the initial simulated measurement components.

233
On the Solution of Some Microwave Problems by an Analog Computer, by D. M. Cyck (EAI Computation Center) and A. Norris (Varian Assoc.); 1958 WESCON CONVENTION RECORD, pt. 1, pp. 70–85.

Three examples of an analog computer method for solving the algebraic equations which arise in microwave theory are presented. The examples are concerned with the bandwidth response characteristics of a stagger-tuned six-cavity klystron amplifier, the bandwidth response characteristics of a multiple-element impedance-matching system, and the propagation characteristics for waveguides partially filled with two dielectric media as a function of frequency, relative dimensions, and properties of the materials. The advantage of the analog computer method of solution is that the computer yields a one-design-parameter family of curves for the frequency range of interest when, as in the problems treated, the time dependence is given by $e^{j\omega t}$. The effect of design parameter changes on the bandwidth response can easily be seen by manually adjusting potentiometer settings on the control panel of the computer.

234
Analogue Computer Determination of Certain Aerodynamic Coefficients, by C. H. Murphy (U. S. Ballistic Res. Labs.); *U. S. Gov. Res. Repts.*, vol. 30, p. 576 (A); December 12, 1958. PB 134635.

The use of a precision electronic analog computer to obtain the coefficients of the differential equation of motion from discrete

measurements of actual motion is described. The Kelley-McShane equations of yawing motion of free flight missiles were fitted and are derived in an appendix.

235
Freezing of Slurry Around Wood and Concrete Piles, by R. F. Scott (U. S. Army Arctic Construction and Frost Effects Lab.); *U. S. Gov. Res. Repts.*, vol. 30, p. 460 (A); November 14, 1958. PB 127112.

Work carried out on an electronic analog computer as part of a program of investigations into heat flow phenomena involving freezing and thawing in soil is described. Since such problems are very complicated mathematically, and only a few exact solutions have been obtained in simple cases, the use of approximate methods of solutions, such as the computer, is mandatory for practical problems. Various checking procedures have been employed to insure the accuracy of the results presented.

236
Operating Experience with West Penn Power Company's Economic Dispatch Computer, by W. R. Hamilton and W. H. Osterle (West Penn Power Co.); *Power Apparatus and Systems*, no. 38 (*AIEE Trans.*, pt. III, vol. 77), pp. 702–707; October, 1958.

Recent operating experience with a computer designed to optimize generator load dispatching for an electric power company is described. The computer solves coordination equations by an analog method. A savings of about $50,000 per 1000 megw of system peak load has been realized through the use of the computer.

E-1: MATHEMATICS—LOGIC—THEORETICAL MATHEMATICS

237
On the Generation and Testing of Random Digits, by H. A. Meyer, L. S. Gephart, and N. L. Rasmussen (Univ. of Florida); *U. S. Gov. Res. Repts.*, vol. 30, p. 398 (A); November 14, 1958. PB 134471.

The historical facts connected with the generation of generally available sets of random digits are outlined and the concepts of randomness and local randomness are defined and discussed. The four standard tests for local randomness proposed by Kendall and Smith are described, and a modification of one of these, the gap test, is proposed. A test for the local randomness of sets of random digits to be used in certain Monte Carlo calculations is also described.

238
Significant Digit Computer Arithmetic, by N. Metropolis and R. L. Ashenhurst (Univ. of Chicago); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 265–267; December, 1958.

The usual floating point arithmetic makes error analysis difficult. An alternative system, called significant digit arithmetic, which offers a means of analyzing floating point calculations more effectively and which also possesses certain advantages from an equipment standpoint is described.

239
A Method for the Reduction of Empirical Multi-Variable Functions, by C. D. Allen (Air Training Link Ltd.); *Computer J.*, vol. 1, pp. 196–200; January, 1959.

A function of several variables, whose values have been empirically determined, is often much more conveniently expressed in the form

$$\phi(x, y, \cdots) = \sum_{\nu=1}^{n} \phi_\nu(x)\psi_\nu(y)$$

where the $\phi_\nu(x)$, $\psi_\nu(y)$, etc., are functions of one variable only. A matrix method, involving the determination of the n greatest eigenvalues of a subsidiary symmetric matrix, is developed for optimizing the functions $\phi_\nu(x)$, etc. n depends on the degree of accuracy required. Finally, rotation by a transformation matrix makes one of the $\phi_\nu(x)$ nearly constant and hence simplifies the results.

240
A New Programming Technique for Rational Fractions, by K. M. Howell (Univ. of Southampton); *Computer J.*, vol. 1, pp. 176–178; January, 1959.

A technique for doing arithmetical computations on rational fractions whose numerator and denominator contain factorials is described. Each factorial is stored in terms of exponents of its prime factors. Multiplication and division is then easily carried out by adding or subtracting these factor exponents. The method is being applied in the construction of a table of $6j$-Wigner coefficients, in the field of nuclear physics.

E-2: MATHEMATICS—LOGIC—SYMBOLIC LOGIC, BOOLEAN ALGEBRA

241
Analysis of Sequential Machines II, by D. D. Aufenkamp (Lockheed Aircraft Corp.); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 299–306; December, 1958.

Mealy's model of a sequential machine is assumed, and a relation of "compatibility" of states is introduced to further the analysis of such machines. In the event that input restrictions exist, it is often possible to effect combinations of states under this relation in addition to those permitted under equivalence of states, a relation previously studied. Compatibility of states is analyzed by an iterative technique, rigorously established, which makes it possible to determine readily connection matrices of simpler "compatible" machines.

242
The Multipurpose Bias Device: Part II, The Efficiency of Logical Elements, by B. Dunham, D. Middleton, J. H. North, J. A. Sliter, and J. W. Weltzien (IBM Corp.); *IBM J. Res. & Dev.*, vol. 3, pp. 46–53; January, 1959.

The efficiency of a logical element can be equated with the set of subfunctions it realizes upon biasing or duplication of inputs. Various classes of elements are considered, and optimum or near-optimum examples are presented. Some related areas of study are suggested.

243
Iterative Combinational Switching Networks—General Design Considerations, by E. J. McCluskey, Jr. (Bell Tel. Labs.); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 285–291; December, 1958.

An iterative network is a combinational switching circuit which consists of a series of identical "cells" or subnetworks; for example, the stages of a parallel binary adder. A formal design method for iterative networks which is similar to the flow table technique for designing sequential circuits is presented.

244
Minimal "Sum of Products of Sums" Expressions of Boolean Functions, by S. Abhynakar (Princeton Univ.); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 268–276; December, 1958.

The problem of economical synthesis of circuits for digital computers leads to the problem of finding Boolean expressions of minimal length equivalent to a Boolean expression f. Previous authors restricted themselves to "sum of products" expressions; dualizing this gives "products of sums." The next more efficient step is to find minimal "sums of products of sums" expressions. The basic concepts of this method are formulated and general theorems are given. All the distinct minimal "sum of products of sums" expressions are obtained for the case in which the cell complex of f consists of two isolated points.

245
Some Properties of Boolean Equations, by N. Rouche (Lovanium Univ., Belgian Congo); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 291–298; December, 1958.

Solubility conditions for a set of Boolean equations are established, first with respect to one variable, then with respect to all variables. By consideration of relations between minimal terms, a simple matrix form is deduced for Boolean equations. Using finite group theory and the properties of the characteristic equation of the matrix, a classification is introduced for Boolean mappings and their iterations, to which corresponds a classification of sequential machines.

246
Folding of Symmetric Functions, by G. P. Weeg (Michigan State Univ.); IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, p. 325 (L); December, 1958.

A necessary condition for the folding of symmetric functions which includes a larger class of subscripts than that considered by Shannon and Caldwell is given.

247
Symbolic Logic and Automatic Computers, by E. Berkeley (Berkeley Enterprises); *Computers and Automation*, vol. 7, pp. 18–20, November, 1958; pp. 28–29, December, 1958; pp. 18–20, 22, 23, January, 1959.

The principles of symbolic logic are reviewed briefly.

E-3: MATHEMATICS—LOGIC—NUMERICAL ANALYSIS

248
Method for Finding the General Solution to an Arbitrary Non-Singular System of Linear Equations Involving $n^3/2$ Multiplications, by J. P. Roth (Princeton Univ. Inst. for

Advanced Study); *U. S. Gov. Res. Repts.*, vol. 30, p. 396 (A); November 14, 1958. PB 127053.

In numerical analysis, it is clear that the order in which algebraic operations are performed affects the number of arithmetic operations that have to be made: for example, if in the expression $a(b_1+\cdots+b_n)$ addition is performed inside the brackets before multiplication by a, n additions and 1 multiplication are performed whereas the opposite order requires n multiplications and n additions. It is shown that if the form in which the general solution to a system of nonsingular linear equations is expressed is suitably modified, the number of arithmetic operations required is cut in half.

249

Mathematics for Digital Computers, Vol. I: Multivariate Interpolation, by W. E. Milne, W. Arntzen, N. Reynolds, and J. Wheelock (Oregon State College); *U. S. Gov. Res. Repts.*, vol. 30, p. 583 (A); December 12, 1958. PB 151200.

The subject of multivariate interpolation and its use in connection with modern stored program digital computers is presented. The discussion is primarily limited to the case of polynomial interpolation, although brief mention of trigonometric interpolation is made. Univariate interpolation is briefly introduced and then bivariate interpolation is discussed in detail. A selective bibliography of 352 references appears at the end of the report.

250

Two Square Root Approximations, by W. G. Wadey (UARC, Remington Rand Corp.); *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 13–14; November, 1958.

The number N^2, whose root is desired, may be expressed as the sum of two numbers A_i^2 and B^2, where A_i^2 is a known square, and $B^2/A_i^2=K$, $(B^2\leq A_i^2)$. Then a first approximation, good to 0.75 per cent, is $N^*=A(1.0075+0.4173K)$. A second approximation, good to 0.064 per cent relative error, is $N^*=A(1.000625+0.485025K-0.7232K^2)$. Error curves, and a set of seven numbers A_i spanning the range (0.01, 1) such that $A_{i+1}^2<2A_1^2$ are given.

251

Error Estimation in Runge-Kutta Procedures, by D. H. Call and R. F. Reeves (Ohio State Univ.); *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 7–8; September, 1958.

The principal criticism of Runge-Kutta methods of differential equation solution has been the inability to estimate the associated truncation error. The idea of reversing directions at each step of the advancing solution and recomputing the previous ordinate is proposed. The difference between this and the originally computed ordinate is utilized as an estimate for the truncation error in odd-order R-K procedures. For even orders, the forward and reverse truncation errors cancel out and the method breaks down. For automatic control of error, truncation is computed at each step, and kept within limits by controlling step-size.

252

A Class of Non-Analytical Iterative Processes, by J. H. Wensley (Computer Dev. Ltd.); *Computer J.*, vol. 1, pp. 163–167; January, 1959.

A method for deriving algorithms for the numerical solution of the equation $f(x)=p$, using no more complex operations than addition, subtraction, doubling, and halving is given. Restrictions on $f(x)$ are that it be nondecreasing in the interval where the solution is known to lie, and that it possess an Addition Theorem of the form $f(s+t)=G[f(s), t]$. The arithmetic is binary, and the solution is developed digit by digit. Examples of division, square (and higher) root, $\log_2$, and inverse cosine are shown.

253

Round-Off Error in the Numerical Solution of the Heat Equation, by J. Douglas, Jr. (Rice Inst.); *J. Assoc. for Computing Mach.*, vol. 6, pp. 48–58; January, 1959.

Round-off error in the numerical solution of the heat equation is analyzed by the backward difference method. This leads to a set of tri-diagonal linear equations to be solved. The propagation of error at each step of the solution is considered and it is demonstrated that if normalized Gaussian elimination is used, the solution is stable and round-off error does not "blow up." If no normalization occurs during the elimination, there is no hope of accurate results for even the first time step.

254

Extraction of Roots by Repeated Subtractions for Digital Computers, by I. Sugai (IBM Corp. Res. Center); *Commun. Assoc. for Computing Mach.*, vol. 1, pp. 6–8, December, 1958; *U. S. Gov. Res. Repts.*, vol. 30, p. 582 (A); December 12, 1958. PB 134512.

An extension to pth roots of the well-known desk-calculator method of extracting square roots by subtraction of successive odd integers (based on the result that the sum of the first n odd integers is n^2) is discussed. For the pth root case, the digits are grouped in units of p decimal digits, and the successive numbers to be subtracted off at each stage are

$$\Delta[a_i^3] = a_i^3 - (a_i - 1)^3.$$

The method, which has been programmed for an IBM 650, takes more memory space than conventional Newton-Raphson methods, but the execution time is cut by 50 per cent for fixed-point arithmetic.

E-4: MATHEMATICS—LOGIC—THEORETICAL LINGUISTICS

255

Three Levels of Linguistic Analysis in Machine Translation, by M. Zarechnak (Georgetown Univ.); *J. Assoc. for Computing Mach.*, vol. 6, pp. 24–32; January, 1959.

A General Analysis Method of Machine Translation that can be characterized in three successive levels which are effected internally by the computer between the input and output stages is developed. The first level concerns the analysis of the individual word, which may be inflected to take many variant grammatical endings. The second level deals with the relations between immediately adjacent words and provides the building blocks out of which sentences may be constructed. The third level solves such problems as locating the nucleus of the noun phrase or verb phrase within the sentence. Sentences are classified into logical types and procedures for translation are designed for each type.

E-5: MATHEMATICS—LOGIC—LINEAR PROGRAMMING

256

What is Linear Programming? by S. Vajda (Admiralty Res. Lab.); *Process Control and Automation*, vol. 6, pp. 15–21; January 1959.

The historical background of linear programming and the mathematical formulation of a linear programming problem are discussed. Several examples of linear programming are given.

J: SUMMARIES AND REVIEWS

257

Survey of Commercial Computers, by N. Macdonald (Computers and Automation); *Computers and Automation*, vol. 7, pp. 8, 10, 12, 13; November, 1958.

Brief descriptions of 25 commercially available analog and digital computers are given. Information on such topics as fields of application, accuracy, components, speed of operation, programming, reliability, and price range is included.

258

Electronic Computers as Tools for Management in the U. S. A.: 1956, by R. H. Gregory (Mass. Inst. Tech.) and H. W. Gearing (Metal Box Co. Ltd.); *Computer J.*, vol. 1, pp. 179–191; January, 1959.

An over-all survey of the state of development and business application of data-processing equipment in the U.S.A. in 1956 is given. Practically every aspect of data processing is touched on, and a series of lessons for prospective users are spelled out. For example, it is pointed out that routines already on punched cards are likely to be readily translatable for machine use and that the cost of organizational preparation for a computer may exceed the cost of the hardware.

259

Ten Years of Computer Development, by the Rt. Hon. Earl of Halsbury (Natl. Res. Dev. Corp.); *Computer J.*, vol. 1, pp. 153–159; January, 1959.

The first ten years of British computers are described beginning with the EDSAC I at Cambridge and the Manchester University Mark I Computer. The faster tempo of U. S. development is ascribed to support from military expenditure, and a readier realization by potential customers of the ultimate value of computers. It is suggested that the British industry cannot compete with present U. S. machines, but should skip immediately to the next generation of computers.

260

Origin and Development of the Chinese Abacus, by S. T. Li; *J. Assoc. for Computing Mach.*, vol. 6, pp. 102–110; January, 1959.

The history and development of the Chinese abacus from its origin about 1100 B.C. in the early Chou Dynasty to the present day is reviewed. In its standard form, the abacus is a bi-quinary device, with a certain logical redundancy in beads, which, however, the special rules for more complicated calculations put to good use. It is pointed out that operations as intricate as matrix inversion can be tackled conveniently on an abacus.

SUBJECT INDEX

AUTHOR INDEX

PGEC News and Notices

Geographical Distribution of PGEC Membership

Because it is felt that the members of the PGEC would be interested in the geographi-distribution of membership, the following listing by IRE Section, as of January 1, 1959, is presented.

Asterisks denote the location of active PGEC Chapters. If you are in an area which does not have access to an existing chapter, and are interested in the formation of one, please contact:

William S. Speer, *Chairman*
PGEC Sectional Activities Committee
Norden Div., United Aircraft Corp.
Gardena, Calif.

Section	*Number of Members*
*Akron (Ohio)	39
Alamogordo-Holloman (N. M.)	15
Albuquerque-Los Alamos (N. M.)	46
Anchorage (Alaska	2
Atlanta (Ga.)	47
*Baltimore (Md.)	113
Bay of Quinte (Ont., Can.)	7
Beaumont-Port Arthur (Tex.)	2
*Binghamton (N. Y.)	92
*Boston (Mass.)	576
Buenos Aires (Argentina)	4
Buffalo-Niagara (N. Y.)	36
Cedar Rapids (Iowa)	24
Central Florida	69
Central Pennsylvania	14
*Chicago (Ill.)	224
China Lake (Calif.)	6
Cincinnati (Ohio)	34
Cleveland (Ohio)	44
Colombia	2
Columbus (Ohio)	15
Connecticut	130
*Dallas (Tex.)	39
*Dayton (Ohio)	74
Denver (Colo.)	30
*Detroit (Mich.)	158
Egypt	1
Elmira-Corning (N. Y.)	10
El Paso (Tex.)	15
Emporium (Pa.)	2
Erie (Pa.)	3
Evansville-Owensboro (Ind.)	6
Florida-West Coast	18
Fort Huachuca (Ariz.)	5
Fort Wayne (Ind.)	10
Fort Worth (Tex.)	39
Hamilton (Ont., Can.)	9
Hawaii	6
*Houston (Tex.)	52
Huntsville (Ala.)	34
Indianapolis (Ind.)	30
Israel	20
Ithaca (N. Y.)	12
Kansas City (Mo.)	32
Little Rock (Ark.)	4
London (Ont., Can.)	4
†Long Island (N. Y.)	366
*Los Angeles (Calif.)	1025
Louisville (Ky.)	12
Lubbock (Tex.)	4
Miami (Fla.)	16
Milwaukee (Wis.)	62
*Montreal (Que., Can.)	34
Newfoundland (Can.)	1
New Orleans (La.)	15
†New York (N. Y.)	805
North Carolina	15
Northern Alberta (Can.)	3
†Northern New Jersey	355
Northwest Florida	3
Oklahoma City (Okla.)	10
Omaha-Lincoln (Nebr.)	9
Ottawa (Ont., Can.)	26
*Philadelphia (Pa.)	527
Phoenix (Ariz.)	73
*Pittsburgh (Pa.)	76
Portland (Ore.)	11
Princeton (N. J.)	53
Quebec (Can.)	9
Regina (Sask., Can.)	2
Rio de Janeiro (Brazil)	1
Rochester (N. Y.)	49
Rome-Utica (N. Y.)	33
Sacramento (Calif.)	13
St. Louis (Mo.)	39
Salt Lake City (Utah)	17
San Antonio-Austin (Tex.)	17
San Diego (Calif.)	88
*San Francisco (Calif.)	401
Schenectady (N. Y.)	42
Seattle (Wash.)	63
Shreveport (La.)	7
South Bend-Mishawaka (Ind.)	9
South Carolina	5
Southern Alberta (Can.)	4
Syracuse (N. Y.)	79
Tokyo (Japan)	48
Toledo (Ohio)	5
Toronto (Ont., Can.)	43
Tucson (Ariz.)	14
Tulsa (Okla.)	23
*Twin Cities (Minn.)	167
Vancouver (B. C., Can.)	18
Virginia	33
*Washington (D. C.)	342
Western Massachusetts	18
Western Michigan	2
Wichita (Kans.)	7
Williamsport (Pa.)	0
Winnipeg (Man., Can.)	4
Total	7257

* Official PGEC Chapters.
† Joint Chapter, three Sections.

Industrial Dynamics to Be Subject of Summer Program

The M.I.T. School of Industrial Management will offer a special two-week program on Industrial Dynamics, August 17–28, 1959. Topics to be covered include mathematical formulation of flows of information, decisions, material, money, capital equipment, and manpower; digital computer simulation studies of the causes of industrial growth and fluctuation; and discussions of future applications to production and distribution, consumer market dynamics, research and development management, and capital equipment industries.

The program is designed particularly for the executive who wishes to evaluate the future importance of new developments in management systems to his corporation. For background information see the article, "Industrial Dynamics—A Major Break Through for Decision Makers," by Prof. Jay W. Forrester of M.I.T., in the July-August issue of the *Harvard Business Review*, 1958. For announcement and application form write the Summer Session Office, Massachusetts Institute of Technology, Cambridge, Mass.

Abstracts Due for Conference on Magnetism and Magnetic Materials

The Fifth Conference on Magnetism and Magnetic Materials will be held in Detroit, Mich., November 16–19, 1959, at the Sheraton-Cardillac Hotel. This conference is sponsored by the American Institute of Electrical Engineers in cooperation with the Office of of Naval Research, the Metallurgical Society of the AIME, the American Physical Society, and the Institute of Radio Engineers. Abstracts should be received by J. E. Goldman, Scientific Laboratory, Ford Motor Co., P.O. Box 2053, Dearborn, Mich., by August 25. Instructions to authors as well as further conference details can be obtained from D. M. Grimes, Department of Electrical Engineering, University of Michigan, Ann Arbor, Mich.

Fourteenth Annual Meeting of the Association for Computing Machinery

The Fourteenth Annual Meeting of the Association for Computing Machinery will be held at the Massachusetts Institute of Technology, Cambridge, Mass., on September 1–3, 1959. Local arrangements will be under the direction of Prof. F. M. Verzuh.

Contributed papers concerned with all phases of analog and digital computation are included on the following topics. Scientific calculations, computer design, information retrieval, theory of automata, numerical analysis, inventory, teamed computers, University applications, programming, applied mathematics, traffic control, coding theory, administration of computers, data processing, united programming efforts,

neurophysiological models, compilers, language translation, paralleled computers, simulation payroll, learning concepts, generators, special devices, management problems, models, operation control, computer communications, computer languages, pattern analysis, computer operating systems, decision making, sorting algorithms, military applications, linear programming, theory of data processing, user groups Monte Carlo techniques, satellite orbits, character reading, medical applications, river control damage assessment, war games, letterwriting logic in design memory units, differential analyzer, iteration, visual display, abaci, errors.

Seminar on Soviet Computers

A Seminar on "Status of Digital Computer and Data Processing Developments in the Soviet Uunion" was held at ONR last November 12. The Proceedings, containing full papers by Professors J. W. Carr, A. J. Perlis, J. E. Robertson, and N. R. Scott, and a discussion session chaired by M. C. Yovits are now available from:

Office of Technical Services
U. S. Dept. of Commerce
Washington 25, D. C.
#PB 151634 Price: $3.00

Please make checks payable to O.T.S., Department of Commerce.

AIEE Letter of Explanation

The following letter has been received by Richard O. Endres, PGEC Chairman, in connection with the Nonmember AIEE legends that appeared under the names of many authors in the *1958 Proceedings of the Western Joint Computer Conference.*

Dear Mr. Endres:

In connection with the Western Joint Computer Conference publication, my attention has been called to the nonmember listings after the names of many authors. This was an unfortunate oversight which came about because it is our standard practice in connection with all *Transactions* papers to list nonmembers after the author's name if he is not a member of AIEE. Nonmember *Transactions* papers require special approval and the reasons why such a paper should be accepted, and our Membership Committee some years ago requested us to print nonmember on all AIEE papers.

This was entirely an oversight on the part of our editors, which came about by not realizing when working on the individual papers that the Western Joint Computer Conference is a jointly sponsored conference, and I am very sorry that it has occurred.

Very truly yours,
Charles S. Rich
Editor and Manager of
Publications, AIEE

Call for Papers for 1959 Eastern Joint Computer Conference

The 1959 Eastern Joint Computer Conference, sponsored by AIEE, ACM, and IRE, will be held at the Statler Hilton Hotel, Boston, Mass., on December 1–3, 1959. Papers on all phases of computing will be appropriate. Present plans call for a single-session conference, and each paper will be limited to a presentation time of 20 minutes followed by a brief discussion period. At the discretion of the program committee, papers of exceptional interest may be allowed a longer period of time for presentation, provided the author makes written request at the time the required abstract and summary are submitted.

Those who wish to present papers should submit four copies of a 100-word abstract and a 1000-word summary by August 15, 1959, to J. H. Felker, Chairman, EJCC Program Committee, Bell Telephone Laboratories, Mountain Avenue, Room 5C-101, Murray Hill, N. J.

An award of $300 will be made for the best paper describing significant work in the computer field.

Summer Session on Finite and Infinite State Machines To Be Offered at M.I.T.

A special summer program in "Finite and Infinite State Machines" will be offered by the Electrical Engineering Department of the Massachusetts Institute of Technology, with the assistance of lecturers from the Department of Mathematics, Linclon Laboratory, and other sources. The program is under the direction of Professors Dean Arden and David Huffman of the Electrical Engineering Department and will consist of a selection of topics in the area of finite and infinite state machines. These topics will range from the theory of sequential switching circuits to the theory of recursive functions. The question of physical realizations of machines will not be discussed. While the amount of background material required will be small, the program will include some rather sophisticated topics, Registrants should have, or be able to acquire quickly, a working familiarity with Boolean algebra. In general the Bachelor of Science degree in mathematics, physics, or electrical engineering will be expected. There will be approximately four hours of lecture each day, divided equally between the morning and afternoon session. No academic credit is offered.

The idea of the description of a logical system in terms of transitions between states appeared in print as early as 1936 in Turing's paper, "On Computable Numbers, with an Application to the Entscheidungs Problem," in the *Proceedings of the London Mathematical Society*, volume 42. Turing considered a finite state machine whose transitions are controlled by and in turn alter the symbols recorded on an arbitrarily long but finite portion of an infinite tape. The basic idea of a state description has since been used by D. A. Huffman ("The Synthesis of Sequential Switching Circuits," *Journal of the Franklin Institute*, 1954) in the solution of some of the problems encountered in the synthesis of sequential machines, and by E. F. Moore ("Gedanken Experiments of Sequential Machines," *Automata Studies*, 1956) in the investigation of the equivalence of automata as determined by external experiments. The concept has also been used by Kleene ("Representation of Events in Nerve Nets and Finite Automata," *Automata Studies*, 1956) in a basic study of the theoretical capabilities of finite state machines.

In general the lecture topics will deal both with the abstract state description of logical machines and the more detailed description as an interconnected collection of basic circuit elements. Many problems concerned with these models remain to be solved and will be indicated in the lectures.

In addition to Professors Arden and Huffman, Professors Edward Arthurs, Samuel Caldwell, and Peter Elias of the Electrical Engineering Department, Professors Marvin Minsky and Hartley Rogers of the Mathematics Department, Dr. Belmont Farley of Lincoln Laboratory, and Frederick Hennie and Leo Jedynak of the Electrical Engineering Department, will lecture.

JOINT COMPUTER COMMITTEE

JCC SENEWS

SCIENCE EDUCATION SUBCOMMITTEE NEWSLETTER

Vol. 2, No. 2 June, 1959

The Western Joint Computer Conference was held in San Francisco, Calif., on March 3–5, and one of the innovations was a special session on industry's role in supporting high-school science programs. It was the first time that a JCC event allowed the students themselves to present reports on their work. The material presented for the most part was interesting; the speakers were articulate and showed a degree of confidence and poise that was quite disarming. Much of the success of this session was due to the efforts of Richard Melville, and his summary is presented in the next section.

Future conferences will see more of this type of activity, and it is highly probable that the Eastern Joint Computer Conference (Boston, Mass., December, 1959) or the next WJCC (San Francisco, May, 1960) will include a session devoted entirely to students, where the papers will be selected on a competitive basis. More will be said about this as the plans become more definite.

In this issue we start publication of short tutorial articles which we feel are suitable for distribution to students. Reprints are available upon request, and a list of other reprints is given at the end of this newsletter. Again, we solicit your cooperation in submitting worthwhile projects or technical articles for student use.

Michael Warshaw, *Chairman*
JCC Science Education Subcommittee
The RAND Corporation
1700 Main Street
Santa Monica, Calif.

Industry's Role in Supporting High-School Science Programs

An introductory statement was presented by J. Paivinen, a member of the WJCC Technical Program Committee. He pointed out that the objective of the meeting was to bring information on cooperative industry-school programs to the attention of interested computer people as well as invited high-school principals from the Bay area (20 out of the 100 invitees attended). Through this meeting the computer industry expressed interest and willingness to cooperate in additional similar programs wherever the need may be evidenced by schools or teachers. It is not the intent to restrict such programs to involve computer technology only, but to participate in broad science programs including a computer portion. In this way, students aiming at careers in science will be exposed to computer techniques as one of their available tools, while potential computer people will have attained a broader grasp of basic technology.

Dr. Paul Hurd, Chairman, Professor of Science Education, Stanford University. Dr. Hurd portrayed the program as illustrating examples of cooperation between industry, professional people, and civil groups working with high schools to enrich the technical exposure of high-school students. He pointed out that the needs of our society have expanded from one in 300 employees in 1900 needing a scientific, engineering, or mathematical training for their positions, to one in 37 needing such training in 1958. The trend, he said, is still increasing for such training. Conversely, schools need the help of industry and civil groups to establish such programs, not only to avoid increasing the teacher's burdens, but to bring to the teachers a knowledge and background in newer areas of technology and science. Examples in the program will point out that the success of cooperative programs depends always on both an interested teacher and interested community members to participate in the extra work. In summary, Dr. Hurd stated that the examples in the program show how a new dimension can be brought to education in which the community itself takes an active part by contributing time, technical skills, and support to enrich high-school and junior high-school programs.

Darryl Littlefield, Physics and Science Teacher, Livermore Union High School. An elective course is offered as part of the curriculum at the high school to cover applications and programming of the IBM 650. The objective of the course is to convey a real appreciation of the power and application of computers in mod-

ern research, engineering, and business. Problems are generally of a nature that would not normally be encountered by high-school students. Mr. Littlefield and Mr. Herman Thomas of the Mathematics Department cooperate in the project.

The course stemmed from Mr. Littlefield's contact with the IBM 650 at the Lawrence Radiation Laboratory (University of California) at Livermore and his consequent interest toward using the community resources to establish programs of real educational growth for his students. Dr. Sidney Fernbach of the Radiation Laboratory acts as consultant for the course, while the Laboratory makes machine time available for the students. The Radiation Laboratory has also made available charts, models, and speakers for the classroom.

Mr. Littlefield mentioned in particular the benefit of a week-long seminar on computer instruction for teachers and engineers, arranged by Dr. Van Atta of Hughes Aircraft Company of Los Angeles.

As another part of the school program, Mr. Littlefield mentioned the construction of a relay binary adder-subtractor by two of his students, using parts donated by the Pacific Telephone and Telegraph and IBM. (*Editor's Note:* We are attempting to obtain a technical description and schematic of this adder, and hope to have it for the next issue.) Dr. LaFrangi of the Radiation Laboratory acted as advisor.

Joanne Watkins, Senior student, Livermore High School. Miss Watkins described her programming to solve the motion of a projectile in vacuum. She showed flow charts, coding sheets, and plots of trajectories for varying initial elevations. She commented that she enjoyed the course because it was new and interesting, and has given her a background applicable even to other computers, together with an appreciation for what jobs can best be done on computers.

Doug McMilin, Junior Student, Livermore High School. Mr. McMilin described a payroll calculation on the IBM 650. He showed the input and output card layouts, flow charts, and programming sheets and described the main points of the calculations. He said that he found personal access to a machine fascinating and enjoyed seeing mathematical principles in practical use. He also considered valuable the training in logical thinking which resulted from working a problem as a whole, with simultaneous attention to detail, while observing the strict rules of computer programming.

Tom Doyen and Ross Harrower, Students, Livermore High School. Messrs. Doyen and Harrower explained the operation of the binary adder they had constructed, including mention of subtraction by complements. A demonstration followed the description.

Dr. Sidney Fernbach, Staff Member, Lawrence Radiation Laboratory at Livermore. Dr. Fernbach described other cooperative programs supported by the Radiation Laboratory. A High School Committee attempts to get students interested in individual projects, providing advisers on such topics as rocketry and nuclear energy.

Speakers are also provided for Science Clubs and classrooms and student tours are arranged at both Berkeley and Livermore Laboratories. Training of high-school teachers in modern physics is arranged through summer and part-time employment where three lectures per week are provided on computers, modern physics, and chemistry.

High-school students have been employed during the summer to convey a working knowledge of laboratory procedures; this summer, participation requiring even *Q* clearance will be arranged.

Finally, general scholarship is encouraged by trying to increase student enrollment in college preparatory courses and by arranging for awards from the community for outstanding students.

Henry Martin, Physics Teacher, Palo Alto High School. Five or six years ago, Mr. Martin had started to seek industry help in guiding the efforts of students who evidenced their eagerness to learn by after-hours and Saturday use of school facilities and the laboratory. However, he encountered little success due to the lack of any concrete program that he and the short-handed science staff had time to generate. This stalemate was broken in 1957 by a consultant from the Joe Berg Foundation, 1712 South Michigan Avenue, Chicago 16, Ill., an organization that volunteers help in establishing an initial relation between industry and schools (*SENEWS*, December, 1958). This resulted in the Palo Alto Science Seminar, which is an entirely locally-conducted program. Fifty-two weekly sessions are conducted throughout the year, with 1 to $1\frac{1}{2}$ hours of each session devoted to a general program given by an industry expert, or to a panel discussion with subsequent group discussions specializing in chemistry, geology, biology, physics, engineering and mathematics. A student joins at about the tenth grade level and will ultimately undertake a project of interest to him; the counselling of a volunteer industry professional will be available. Industry also provides equipment, speakers, demonstrations, and sometimes facilities where the student often works in proximity to an engineer or a scientist. Advantages to the students include familiarization with work conditions and opportunities, opportunity for individual creative study (since projects are not group efforts), personal satisfaction of a hobby with prestige value, and the opportunity to seek and earn scholarships.

Advantages to industry include an early encouraging hand to potential Ph.D. scientists, an opportunity to demonstrate to a broad slice of the community that scientists are normal human beings with families and a sense of humor, and an opportunity to contribute to the community and the future of our country.

In closing, Mr. Martin recommended other programs for industry's consideration:

1) Participation of students and of teachers at professional dinners and educational programs arranged by industry.
2) The assignment to some member of a company management team of the specific responsibility of support and cooperation in educational affairs.
3) The opportunity for students on field trips to spend sufficient time with the engineers and scientists to gain some insight into the significance of the projects or laboratories visited.
4) Summer jobs for students which will increase their knowledge of industry practices and expectations.
5) Screening and testing programs to select outstanding students to receive substantial scholarships.
6) Summer schools at local universities so that students can see what they will be up against in the future.
7) Public competitions for scholarship so that recognition could be given in a manner comparable to school athletic letters and Father's Club dinners for athletes.

Larry Hubbart, Student, Palo Alto High School. Mr. Hubbart described his construction of a test stand as well as his subsequent experiments in measuring the lift of rotating airfoils (similar to an inverted pie plate). An engineer from Hiller Helicopter Company acts as adviser; his contribution was described as helping to suggest directions of investigation to pursue as well as to maintain morale when the project appears to bog down. The test stand consisted of a counter-balanced scale with a drive motor and a photocell rpm counter. Airfoils of varying shapes and surface textures, and with added ducts, have been measured. Mr. Hubbart described the project as representing a success "even if it never leaves the ground" due to the experience it provided in how to conduct an experiment and how to present the results.

Mike Macauley, Student, Palo Alto High School. Mr. Macauley is conducting an experiment in combatting muscle fatigue by injected chemical solutions. He anesthetizes a rabbit and causes muscle contractions by mild electric shock until exhaustion occurs. Injection of a mild hydrogen peroxide solution results in acceleration of the recovery: recovery periods range from 3 minutes for a 3 per cent solution to 0.3 second for 15 per cent. Hydrogen peroxide was chosen due to the safe decomposition products; procedures are followed to safeguard against bubble formation in the blood.

Mr. Macauley describes that the insights gained into medicine and medical research have reinforced his interest in entering the field of medicine.

Wallace Burton, Engineer, R-S Electronics, Palo Alto. The YMCA Men's Club, in seeking to further encourage science students and to provide recognition for work done in the Palo Alto Science Seminar, chose to present an annual Palo Alto Industry-Youth Science Show. The show presents awards to outstanding student science projects and also presents an opportunity for exhibits by local industrial firms.

Announcements of the show with award categories are sent to all the high schools in the Palo Alto district and a participation of about 10 per cent of the students is experienced (the only other such event, the Bay Area Science Fair, accommodates only 1 per cent of the Palo Alto students). Three categories of participation are provided: physical sciences, biological sciences, and technical reports (given orally in competition). About one-half of the industrial firms contacted responded favorably to requests for financial support and industrial exhibits. In the first year, industry met $650 of the $850 expenditures; this year the show will be self-supporting, and it is expected to eliminate the club deficit.

The Science Show provides wide recognition to students through substantial adult attendance as well as the interest and awareness of the industry people who work as judges and exhibitors for the show. The Men's Club provides the planning and organization for the show so that the participation of the various schools can be drawn together into an integrated plan; this has been found necessary to obtain effective support from industry.

Mr. Burton suggests that even smaller communities can establish Science Shows. Community members are sure to lend help if given an opportunity while local commercial businessmen would undoubtedly help support the expenses.

Staff Member, System Development Corp., Los Angeles. A member of the audience described educational support activities performed by the staff at System Development Corp. At the junior high-school level, package lectures are available in data-processing, automatic feedback, and binary and octal notation in programming, etc. For the high-school level, talks are available in concepts of programming, and environmental simulation in air defense, etc. City colleges offer computer design philosophies, logic design, and computer system design, etc. A symposium has been held for teachers on "Implications of the Computer Age for Teaching Mathematics," while other lectures and special courses on programming and advanced mathematics for high-school students have also been provided. (*Editor's Note:* For further information contact Warren Pelton, System Development Corp., 2500 Colorado Ave., Santa Monica, Calif.)

JCC Science Education Literature

Projects

[1] Decimal to Binary Converter No. 1 (Schematic only).
[2] Decimal to Binary Converter No. 2 (Schematic only).

Articles

[1] M. Warshaw, "How Electronic Computers Work."
[2] G. E. Forsythe, "Bibliography on High-School Mathematics Education."
[3] W. G. Schmidt, "Boolean Algebra and the Digital Computer."

HOW ELECTRONIC COMPUTERS WORK

The purpose of this article is to describe briefly the modern, high-speed electronic computer. The subject is a rather complicated one, so a complete and rigorous treatment is out of the question in these few pages. We would like, however, to present a few of the fundamentals and dispel the mystery which motivates people to attach to the computer such names as "electronic brains," etc.

It is first necessary to distinguish between two major classes of computers, those known as *analog* computers and those known as *digital* computers. The analog computer is a conceptually simpler machine, so we will deal with this type first.

Analog computers do not deal with numbers *per se*. They deal, instead, with an *analog* of an *actual physical system*. To give an immediate example to illustrate this statement. Suppose we wish to determine the behavior of a missile in flight, and that we have in our possession all the mathematical equations which describe its behavior. Furthermore, we are told that this problem is to be solved on an electronic analog computer.

The computer is first *connected* in such a fashion that the mathematical operations called for in the equations of the problem can be carried out. But, within the computer, the *physical quantities* in which we are interested, such as missile velocity and altitude, fuel supply, various pressures, etc., are represented by *voltages*. These voltages are the analogs of the actual physical quantities, and it is these voltages that are operated upon by the computing circuits within the machine. The circuits within the machine are electron tube circuits that are capable of accepting, as their inputs, voltages which represent physical quantities, and producing, at their outputs, other voltages which represent, say, the sum or difference of these quantities. There are also circuits that carry out the operation of multiplication, division, and the calculus, thus permitting quite complex physical situations to be evaluated. It is, of course, up to the operator of the computer to see that these circuits are connected in such a manner that the actual physical problem is represented.

So far, we have not had to deal with numbers at all. The mathematical equations that represented the physical situation were wired into the machine, and various voltages that represent physical quantities were manipulated by these circuits so as to produce some results. But we might well ask at this point, how do we introduce the constants and initial conditions of the problem; these, most certainly, will come to us as numbers. And what is more important, how do we get the final solutions out of the machine, which must be numbers if they are to be of any use?

This brings us to one of the prime concepts of analog computing, which also happens to be its greatest shortcoming. This is the concept of *measurement*. In order to insert the initial conditions and constants *into* the machine, and in order to get a solution *out* of the machine, we must measure some voltage. It is this measurement process which introduces most of the errors in analog computing.

Fig. 1 should help at this point. The problem is to evaluate $Ax+B$, where A and B are constants and x is some independent variable. The two boxes shown are electronic devices capable of forming the product and sum of the two voltages presented at their respective inputs. Therein lies our first problem; if these voltages are to be faithful analogs of the actual quantities with which we are dealing, we must determine the value of the voltage quite accurately. For instance, if the constant A is 23, and if we assign a one-to-one correspondence between the value of the constant and its analog voltage, then we must apply exactly 23 volts at the input terminal marked A. If we carry out this process at all three of the input terminals, namely, A, B, and x, a voltage should appear at the output terminal which is the voltage analog of the desired solution. And therefore, once again, a measurement must be performed to transform this voltage into a numerical solution.

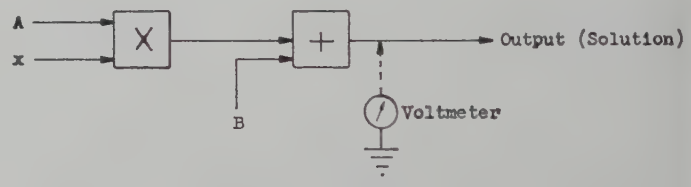


Fig. 1—$Ax+B$ = Solution.

Even if we assume perfect computing elements, *i.e.*, the multiplication and addition are done without introducing any further error, the inherent limitation of this type of computing is our inability to perform sufficiently accurate measurements on the input and output terminals. Accuracy to four or five significant figures is about all that the state of the electronic art will permit, and if errors introduced by the electronic computing circuits are taken into account, this accuracy in many instances drops to three significant figures.

Fig. 2 might help tie together all that we have said about analog computers. We see here where the measurements must take place, and how the mathematical equations actually determine the structure of the computing circuits. For a different problem, the computer is essentially rewired to suit the new equations.

This rewiring and the voltage measurements can usually be done with the greatest of ease, and it is this fact which makes the analog computer such a useful device. Very complex problems can be set up quickly, and the input conditions can be varied easily at the

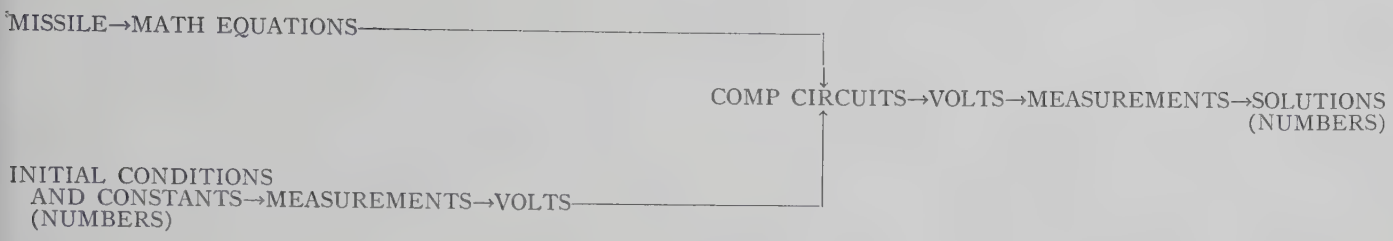


Fig. 2—Flow diagram of an electronic analog computer.

desire of the operator. If extreme accuracy is not important, this is truly the computer for the scientist and engineer. He can try out hunches, modify his equations and constants with little or no trouble, and thus arrive at very good approximate solutions in very little time.

Now let us leave the field of analog computors and move on to the discussion of digital computers. We hope to be able to deal with the digital computer in considerably more detail; it is a more complicated device, and probably a more important one, having become a major industry in this country in the space of very few years. This preferential treatment is also justified by the fact that digital computers have created a tremendous new demand for engineers and mathematicians.

A digital computer deals directly with numbers, and in a manner identical to the way that humans deal with numbers. This is probably the most important point to stress regarding digital computers. It does not, as does the analog computer, put all its eggs in one basket by representing a number by a single voltage. Instead, if we desire to represent the number, say 4096, and we have a computer that works in the decimal system, we will supply four voltages, each voltage representing a single digit in this array of four digits.

If a digital computer is called upon to add numbers, it does as we would do, adding the numbers in the least significant column, determining whether or not there is a carry to the next to least significant column, and then performing another addition on this column. This process continues until the most significant column has been summed, and the addition is now complete.

By adopting this policy, the measurement problem that plagued us in the analog computer has very nearly disappeared. It is necessary for us now only to distinguish the state of any one column. In the decimal system this means we must be able to distinguish among ten allowable states, *i.e.*, the numbers zero through nine. For those computers that work in the binary number system, we only need distinguish between the two states, zero and one. This is a far cry from the analog system, where, if we desired tenth per cent accuracy, we would be called upon to distinguish between 1000 possible values of voltage on the same terminal.

An abacus constitutes an example of a digital computer taken from everyday life (or, at least, everyday oriental life). Here, beads are arranged in columns in a fashion resembling the decimal system. The computation, say, for addition, starts at the least significant end of the board, sums this column, generates the proper carry into the next column of beads, and proceeds until all significant columns have been summed. The question of measurement never comes up as long as we have the small amount of eyesight or tactile sense necessary to ascertain whether or not a particular bead has been moved into play. If we now desire a solution with greater precision, it is necessary only to build an abacus of greater length, *i.e.*, one that is capable of handling longer numbers.

Something that more closely approximates the electronic digital computer is shown in Fig. 3. The pencil and paper serve as a storage media as well as an input/output device. All arithmetic calculations are performed by the desk calculator, and whole operation is under the control of a human being. If we replace these elements with devices that operate at a higher speed, and particularly, if we eliminate the human, we would have an automatic digital computer.

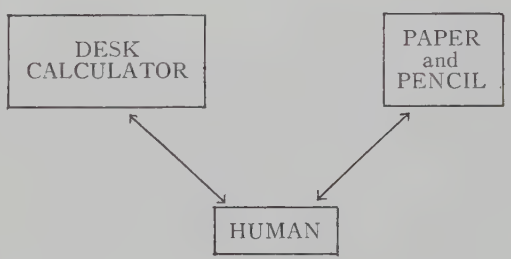


Fig. 3.

Fig. 4 shows those parts that are common to all digital computers. We will discuss each element separately before combining them into a working system.

On the far right is a block that represents the input/output equipment. It is through this device that all data and instructions to the machine are entered. Most computers today use either punched cards of the IBM or Remington-Rand variety, or punched paper tape as the input media. In the case of punched cards, the cards are loaded into a card reader whose output is directed to the storage system of the computer. This process takes place at the rate of approximately 40 ten-digit numbers per second.

Once the machine has been loaded, the computation can proceed. All arithmetic operations are carried out

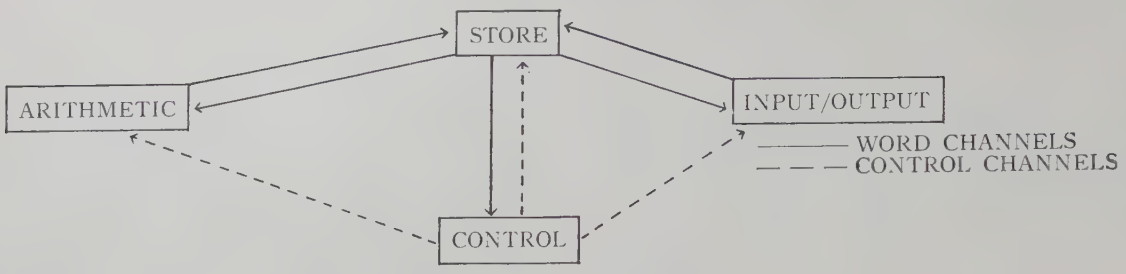


Fig. 4.

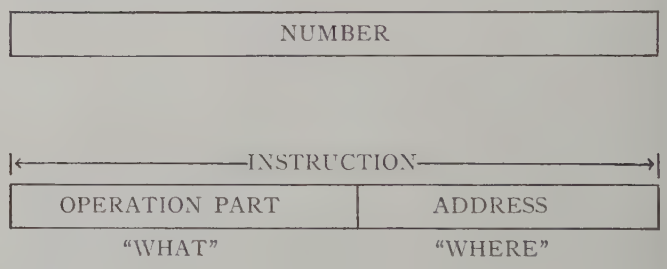
in the arithmetic section of the computer, and a computational program consists of a continuous interchange of data between the store and the arithmetic section. Since this interchange of data must be guided toward some useful goal, and since this was the main task of the human in our last figure, an electronic control section is incorporated to direct the progress of the problem.

One major question needs consideration at this point; namely, how does the computer know what problem to solve? You will remember in the case of the analog computer that the computation elements were actually reconnected in a different manner for each problem. This is *not* the case with the digital computer. The arithmetic section is permanently wired together and is capable of performing the ordinary operations of arithmetic upon receiving an instruction to do so from the control. This still does not answer the question of how the computer knows what problem to solve; to answer it we must consider the contents of the store in considerably more detail.

Remember we are dealing with a digital computer, and that numbers are no longer represented as a single voltage, but rather by a *set* of symbols that are arranged in a particular sequence. It is true, however, that each symbol is usually represented in the machine as a voltage, but since we must only distinguish at most ten and usually only two values (depending on the number system) this is not considered a measurement problem.

So much for the reiteration of the digital concept. Let us continue with the description of the computer store. The store is partitioned into definite cells and associated with each of these cells is what is known as an *address*. This is identical in concept with house addresses in a postal zone, if for no other reason than to permit the postman to deliver the main in a consistent manner. Now, into each cell in the store we place, by means of the input equipment, what is known as a *word*. The temptation to use the word *number* instead of *word* is considerable, as we have been dealing with numbers up to now. However, the contents of a storage cell is not always a number; *i.e.*, it is not always a number in the sense that it is data for use in some computation.

Words can have two forms, and these forms are shown jn Fig. 5. At the top, we see that a word can, indeed, be just an ordinary number that is to be transferred to the arithmetic section at some later date and used in computation. At the bottom, however, we have something new. It is called an *instruction*, and is used by the control section to determine the interplay between the store, the arithmetic section, and the input/output equipment. From all external appearances, an instruction is just a number that has been partitioned into two parts. When this instruction is sensed by the control, it makes the following interpretation: it finds in the left-hand part of the instruction a code number that tells it what to do next. This is the operation part of the instruction. Some typical operations might be add, multiply, or print out. Every operation that the computer is capable of performing has a unique code symbol, and this code symbol is always to be found in the left-hand end of an instruction, where it can be interpreted by the control.



WORDS

Fig. 5.

So far we have stored data and a sequence of instructions in the store, and, by looking at the left-hand part of an instruction, the control knows what operation to perform next. What we have failed to do is tell the control where in the store it is to get the data on which to perform the specified operation. This is the purpose of the right-hand part of the instruction. The right-hand part specifies the *address* of the data that is to be used in carrying out the operation. Remember that each cell in the store has a unique address associated with it, and it is this address that is specified in the right-hand part of the instruction. When this concept is encountered for the first time, it usually causes confusion. The right-hand part of an instruction *does not* contain the data

Problem: To compute $Ax+B$, and print out both x and the solution.

Location in store	Instruction		Comments
	Operation	Address	
01	Output the word in location.	09	Print out x.
02	Take the word in location and place it in the arithmetic.	10	Transfer the constant A to arithmetic section.
03	Take the word in location, multiply it in the arithmetic by the number found there and store the product in the arithmetic section.	09	Compute the product Ax.
04	Take the word in location and add it to the number in the arithmetic section. Store the sum in the arithmetic section.	11	Compute the sum $Ax+B$.
05	Store the word in the arithmetic in location.	09	Transfer $Ax+B$ to the store.
06	Output the word in location.	09	Print out $Ax+B$.
07	Halt.		
08			
09	x		Store the variable x.
10	A		Store the constant A.
11	B		Store the constant B.

Fig. 6.

proper, only the address in the store where the data can be found.

If we refer back to Fig. 4, we can discuss the information flow paths with more intelligence, and it may serve to further illustrate what we have just been talking about. The solid lines represent the paths that words may take as they travel around the machine. Words that travel from the store to the control are instructions. Words that travel between the store and the arithmetic section or input-output equipment are usually data, but this is not always the case as will be mentioned later. The dashed lines are the control signals issued by the control section after it has interpreted some instruction and is carrying out this instruction.

Lastly, let us take the same simple problem that we solved with the analog computer and show how this problem might be solved with a digital computer. Fig. 6 shows the list of instructions that must be written by a computer programmer to enable the computer to deal with this problem. In other words, this is the list of instructions that must be written by the *programmer* so the computer will know what problem to solve. The left-most column shows where, in the computer's store, these instructions are to be placed. Note that in this simple problem we use only eleven storage cells. This is far less than the four thousand or more cells usually available to the programmer. Opposite each location in the store we see the word that is stored in that particular location. Cells 01 through 07 contain instructions, *i.e.*, an operation and a corresponding address. Locations 09 through 11 contain actual numbers that will be used in the process of computation.

Unless told differently, the control will always start at location 01 and proceed in sequence through cells in the store. With that added bit of information, we are ready to start.

The instruction located in cell 01 tells the control to go to cell 09 and print out whatever number is stored there.

Cell 09 contains x, so x is printed. The instruction in 02 says, "go to cell 10, get the number stored there (the constant A) and transfer it to the arithmetic unit."

The entire process will not be described here, and the reader should be able to follow each instruction by himself.

It should now be obvious that digital computers are really very stupid devices. They do nothing that the programmer has not already written into the machine; all phrases such as "electronic brains" are serious misnomers, and their use should be discouraged. There are, however, some extremely important subtleties in the digital computer that we would like to mention in passing. With proper programming, a computer can be made to alter its own program. By transferring an *instruction* to the arithmetic section, and modifying it by some arithmetic operation, the machine can itself form a new instruction for later use. The computer can also make decisions; such as, which of two numbers is the larger or whether or not a number is positive or negative. On the basis of such a decision, it branches to one of two parts of the program and is thus capable of determining its future course of action depending on the present state of the problem. Remember, however, that all of these "decision making" operations are put into the problem by the programmer.

In conclusion, something should be said about the speed and cost of such computers. Some modern high-speed computers are capable of doing operations at the average rate of 300,000 per second. It takes between 1 and 2 millionths of a second to transfer information between the store and the arithmetic section, with the balance of time being taken up performing the actual operation.

However, one must pay the piper for all this ultra-high speed, or if not the piper, at least some manufacturer of digital computers. Machines can now be purchased for as little as $50,000, but for machines that have the speeds mentioned above, the price will probably be closer to $3,000,000; or, if one prefers the easy rental plan, the fee is on the order of $50,000 per month.

This article has been written for the purpose of acquainting the reader with the most basic elements of electronic computers. For those interested in further information, a bibliography is provided. The books by Murphy, Bowden, and Berkeley are the most elementary, and that by Phister is the most advanced.

Bibliography

Digital Computers

E. C. Berkeley, "Giant Brains," John Wiley and Sons, Inc., New York, N. Y.; 1949.

Staff of Engineering Research Associates, Inc., "High-Speed Computing Devices," McGraw-Hill Book Co., Inc., New York, N. Y.; 1950.

B. V. Bowden, "Faster Than Thought" Sir Isaac Pitman and Sons, London, Eng.; 1953.

R. K. Richards, "Arithmetic Operations in Digital Computers," D. Van Nostrand Co., Inc., New York, N. Y.; 1955.

D. D. McCracken, "Digital Computer Programming," John Wiley and Sons, Inc., New York, N. Y.; 1957.

J. T. Culbertson, "Mathematics and Logic for Digital Devices," D. Van Nostrand Co., Inc., New York, N. Y.; 1958.

J. S. Murphy, "Basics of Digital Computers," John F. Rider, Inc., New York, N. Y.; 1958.

M. Phister, "Logical Design of Digital Computers," John Wiley and Sons, Inc., New York, N. Y.; 1958.

Analog Computers

W. W. Soroka, "Analog Methods in Computation and Simulation," McGraw-Hill Book Co., Inc., New York, N. Y.; 1954.

H. M. Paynton, "A Palimpsest on the Electronic Analog Art," George A. Philbrick Research, Inc., 230 Congress St., Boston, Mass.; 1955.

G. A. Korn and T. M. Korn, "Electronic Analog Computers," McGraw-Hill Book Co., Inc., New York, N. Y.; 1956.

LED IN STACKS
30 Nov 64F

INFORMATION FOR AUTHORS

IRE TRANSACTIONS ON ELECTRONIC COMPUTERS is published quarterly, in March, June, September, and December, with a distribution of over 8500 copies, largely to engineers, logicians, and supervisors in the computer field. Its scope includes the design, theory, and practice of electronic computers and data-processing machines, digital and analog, and parts of certain related disciplines such as switching theory and pulse circuits.

If a paper of widespread interest beyond the computer field is submitted, it will be recommended to the Editor of PROCEEDINGS OF THE IRE for publication. If our reviewers feel that a paper should be submitted to a different IRE TRANSACTIONS, we will so recommend to the author.

Publication time in IRE TRANSACTIONS ON ELECTRONIC COMPUTERS, from receipt of the original manuscript to mailing of the issue, is normally in excess of five months, but can be made as little as 3½ months if the occasion demands and the manuscript is carefully prepared.

To avoid delay, please be guided by the following suggestions:

A. Process for Submission of a Technical Paper

1) Send to the Editor three copies of your manuscript, each copy complete with illustrations. (For Letters to the Editor, two copies will do.)

2) Enclose originals for the illustrations, in the style described below. Alternatively, be ready to send the originals immediately upon acceptance of the paper.

3) Enclose a separate sheet giving your preferred address for correspondence and return of proofs.

4) Enclose a technical biography and photograph of each author, or be ready to supply these upon acceptance of the paper. For biography style, see any IRE journal.

5) If the manuscript has been presented, published, or submitted for publication elsewhere, please so inform the Editor. Our primary objective is to publish technical material not available elsewhere, but on occasion we publish papers of unusual merit that have appeared or will appear before other audiences.

B. Style for Manuscript

1) Typewrite, double or 1½ space; use one side of sheet only. (Good office-duplicated copies are acceptable.)

2) Provide an informative 100- to 250-word summary (abstract) at the head of the manuscript. It will appera with the paper and also separately in PROCEEDINGS OF THE IRE.

3) Provide a separate double-spaced sheet listing all footnotes, beginning with "*Manuscript received by the PGEC _____," and "(Affiliation of author)," and continuing with numbered references. Acknowledgment of financial support is often placed at the end of the asterisk footnote.

4) Give complete references, insofar as possible. See footnotes in this or previous issues for examples of IRE style. References will be printed as footnotes in the column where first mentioned.

5) You may choose to provide a "Bibliography" at the end of the paper, with items referred to by a numeral in square brackets, *e.g.*, [12], to supplement or supplant footnote references.

6) Provide a separate sheet listing all figure captions, in proper style for the typesetter, *e.g.*: "Fig. 1—Example of a disjoint and distraught manifold."

C. Style for Illustrations

1) Originals for illustrations should be sharp, noise-free, and of good contrast. We regret that we cannot provide drafting or art service.

2) Line drawings should be in India ink on drafting cloth, paper, or board. Use 8½×11 inch size sheets if possible, to simplify handling of the manuscript.

3) On graphs, preferably show only the coordinate axes, or at most only the major grid lines, to avoid a dense, hard-to-read result.

4) All lettering should be large enough to permit legible reduction of the figure to column width, perhaps as much as 4:1.

5) Photographs should be glossy prints, of good contrast and gradation, and any reasonable size.

6) Number each original on the back, or at the bottom of the front.

7) Note item B-6 above. Captions lettered on figures will be blocked out in reproduction, in favor of typeset captions.

Mail all manuscripts to:

Dr. Howard E. Tompkins, *Editor*
IRE TRANSACTIONS ON ELECTRONIC COMPUTERS,
The Moore School of E.E., University of Pennsylvania
200 South 33rd Street
Philadelphia 4, Pa.